

## THE PANCHROMATIC HUBBLE ANDROMEDA TREASURY XV. THE BEAST: BAYESIAN EXTINCTION AND STELLAR TOOL

KARL D. GORDON<sup>1,2</sup>, MORGAN FOUESNEAU<sup>3</sup>, HEDDY ARAB<sup>1</sup>, KIRILL TCHERNYSHYOV<sup>4</sup>, DANIEL R. WEISZ<sup>6,7</sup>,  
 JULIANNE J. DALCANTON<sup>6</sup>, BENJAMIN F. WILLIAMS<sup>6</sup>, ERIC F. BELL<sup>14</sup>, LUCIANA BIANCHI<sup>4</sup>, MARTHA BOYER<sup>11,12</sup>,  
 YUMI CHOI<sup>6</sup>, ANDREW DOLPHIN<sup>8</sup>, LÉO GIRARDI<sup>10</sup>, DAVID W. HOGG<sup>9,3</sup>, JASON S. KALIRAI<sup>1</sup>, MARIA KAPALA<sup>3</sup>,  
 ALEXIA R. LEWIS<sup>6</sup>, HANS-WALTER RIX<sup>3</sup>, KARIN SANDSTROM<sup>13</sup>, & EVAN D. SKILLMAN<sup>15</sup>

*Accepted for publication in the ApJ.*

### ABSTRACT

We present the Bayesian Extinction And Stellar Tool (BEAST), a probabilistic approach to modeling the dust extinguished photometric spectral energy distribution of an individual star while accounting for observational uncertainties common to large resolved star surveys. Given a set of photometric measurements and an observational uncertainty model, the BEAST infers the physical properties of the stellar source using stellar evolution and atmosphere models and constrains the line of sight extinction using a newly developed mixture model that encompasses the full range of dust extinction curves seen in the Local Group. The BEAST is specifically formulated for use with large multi-band surveys of resolved stellar populations. Our approach accounts for measurement uncertainties and any covariance between them due to stellar crowding (both systematic biases and uncertainties in the bias) and absolute flux calibration, thereby incorporating the full information content of the measurement. We illustrate the accuracy and precision possible with the BEAST using data from the Panchromatic Hubble Andromeda Treasury. While the BEAST has been developed for this survey, it can be easily applied to similar existing and planned resolved star surveys.

*Subject headings:* dust, extinction – galaxies: individual (M31) – methods: data analysis – methods: statistical – stars: fundamental parameters

### 1. INTRODUCTION

The ability to resolve the stellar content of numerous galaxies has revolutionized our understanding of galaxy formation and evolution. The color and luminosity of an individual star encodes information about the star's intrinsic mass, age, and metallicity and further illuminates the intervening dust, revealing the composition, grain size distribution, and column density of the obscuring material. Decoding this information provides new constraints on the mass assembly and chemical history of a galaxy and on the detailed study of a galaxy's interstellar medium structure and dust composition. The high

information content contained in resolved stars has motivated extensive surveys that have cataloged hundreds of millions of stars in the Milky Way and Local Volume (e.g., Zaritsky et al. 1997; Stoughton et al. 2002; Holtzman et al. 2006; McConnachie et al. 2009; Dalcanton et al. 2009, 2012a; Sabbi et al. 2013; Williams et al. 2014). This number will increase at least ten-fold with the next generation of dedicated programs (e.g., LSST, Gaia, PAN-STARRS, etc.).

With the wealth of resolved star data in external galaxies, both available and forthcoming, it is essential to develop robust methods that can fully exploit observables to infer well-characterized physical quantities. Here we concentrate on the observed spectral energy distributions (SED) of single stars, which reflects three different aspects: stellar physics, intervening dust, and observational effects. Individual stars emit photons over a wide range of wavelengths. A portion of these photons are then removed by dust through absorption and scattering out of the line-of-sight, extinguishing the star's intrinsic spectrum. Finally, the precision and accuracy of the measured flux is modulated by the finite number of observed photons, the contrast between the target and local background, the wavelength sensitivity of the selected instrument (e.g., filter sensitivity in the case of photometric observations), and the accuracy with which the flux from neighboring stars can be subtracted. Therefore, recovering the intrinsic properties of the star and the intervening dust requires modeling of both physical and observational effects.

Historically, stellar SED fitting techniques have taken simplified approaches that make this analysis more tractable, especially for observations of sources in our Galaxy. These compromises often involve focusing on

<sup>1</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA; kgordon@stsci.edu

<sup>2</sup> Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium

<sup>3</sup> Max Planck Institute for Astronomy, Koenigstuhl 17, 69117 Heidelberg, Germany

<sup>4</sup> Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA

<sup>6</sup> Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195, USA

<sup>7</sup> Hubble Fellow

<sup>8</sup> Raytheon Company, Tucson, AZ, 85734, USA

<sup>9</sup> Center for Cosmology and Particle Physics, New York University, 4 Washington Place, New York, NY 10003, USA

<sup>10</sup> Osservatorio Astronomico di Padova-INAf, Vicolo dell'Osservatorio 5, I-35122 Padova, Italy

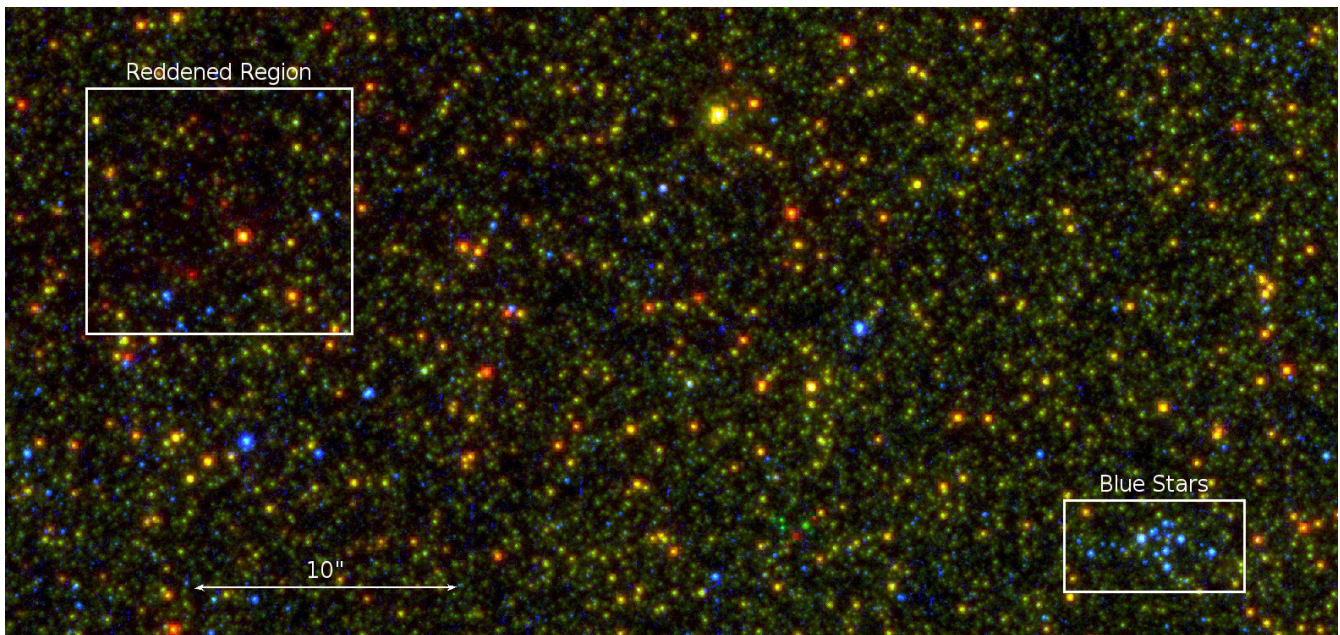
<sup>11</sup> Observational Cosmology Lab, Code 665, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

<sup>12</sup> Oak Ridge Associated Universities (ORAU), Oak Ridge, TN 37831, USA

<sup>13</sup> Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ, 85721, USA

<sup>14</sup> Department of Astronomy, University of Michigan, 500 Church Street, Ann Arbor, MI 48109, USA

<sup>15</sup> Minnesota Institute for Astrophysics, University of Minnesota, Minneapolis, MN 55455, USA



**Figure 1.** A small region from the PHAT survey of M31 in the 10 kpc star forming ring is shown using the F336W (blue), F814W (green), and F160W (red) band images. The entire region shows strong crowding with sources very near each other. In addition, the diverse range of environments is illustrated by the two marked regions, one showing a region that suffers from dust reddening (red sources and fewer sources overall) and the other a blue cluster of stars.

singular science goals such as constraining only stellar metallicities, e.g., the ‘Ultraviolet excess technique’ (e.g., Wallerstein & Carlson 1960), or removing line of sight extinction effects, e.g., the ‘Q’ parameter (e.g., Johnson & Morgan 1953). Additionally, the characterization of uncertainties is often limited only to photon noise and absolute flux calibration uncertainties, motivating the use of conventional  $\chi^2$  fitting techniques. However these approximations can strongly compromise the results from resolved surveys of external galaxies where they are not valid.

For the most part, such techniques have provided insightful astrophysical results using various Galactic and extragalactic datasets (Bianchi et al. 2001; Romaniello et al. 2002a; Zaritsky et al. 2004; Berry et al. 2012; Schlafly & Finkbeiner 2011; Bianchi et al. 2012b). However, these methods have a number of shortcomings that hinder their ability to completely exploit the information content of observations, including the inability to accurately model data in moderate and low signal-to-noise regimes or to incorporate full accounting of the observational uncertainties. Thus, in the era of large, deep, and expensive surveys, we should not limit ourselves to sub-optimal fitting methods to analyze resulting datasets.

In the past decade, the introduction of probabilistic stellar SED fitting techniques has led to significant improvements (Maíz-Apellániz 2004; Fouesneau & Lançon 2010; Bailer-Jones 2011; Bianchi et al. 2012a,b; Da Rio & Robberto 2012; Green et al. 2014; Schönrich & Bergemann 2014; Ness et al. 2015). These approaches are designed to better capture crucial inter-parameter degeneracies and provide flexible frameworks for including different models of stellar evolution, stellar atmospheres, and extinction laws. However, despite the marked improvement in techniques, the current generation of stellar SED fitting codes are still not optimal for analyzing datasets generated by the current and next-generation

resolved star surveys of external galaxies.

One limitation of existing SED fitting codes is in their treatment of dust extinction. Current and new datasets cover diverse ranges in galactic environments showing a range of dust content, necessitating a comprehensive model of interstellar dust extinction to correctly constrain both dust and stellar parameters. A second limitation of current SED fitting codes is in their treatment of source “crowding” wherein the measured flux of a source is affected by the presence of nearby sources. As a result, recovered stellar fluxes are affected by systematic biases and uncertainties due to local crowding that are beyond those expected for random photometric errors due to photon counts (Stetson & Harris 1988). Such systematics can and usually do dominate the photometric error budget for most stars in extragalactic surveys. These effects can be particularly large in some of the most scientifically interesting regions such as stellar clusters and star forming complexes.

Large datasets of resolved stars in external galaxies contain a wealth of information near a given survey’s detection limit, where crowding often induces strong correlations in the measurement uncertainties and becomes the dominant source of uncertainty. Fig. 1 gives an example image from the Panchromatic Hubble Andromeda Treasury (PHAT; Dalcanton et al. 2012b) program that shows strong crowding and a range of dust extinctions. Failure to account for these effects can lead to systematically incorrect inferences about parameters and an inability to fully capture the complete information content of the observations, undermining our ability to correctly interpret large surveys.

Finally, survey observations cover a wide range in stellar spectral types and dust extinctions, implying that not all stars will be detected at a high degree of significance in all filters. For example, the ultraviolet (UV) flux of luminous cool asymptotic giant branch stars can

easily be fainter than the observational limit of the UV bands in a survey. On the other hand, a  $\sim 5 M_{\odot}$  main sequence star may be recovered with moderate signal-to-noise in the UV, but have low signal-to-noise ( $< 3\sigma$ ) in the near-infrared. Moreover even low signal-to-noise detections have important information about the SED and therefore should **be** included when constraining the stellar and dust parameters of a survey.

Our goal is to devise an approach for recovering the intrinsic physical properties of a star (e.g., mass, age, metallicity, etc.) and the intervening dust (e.g., composition, column density, and size distribution) while simultaneously including robust uncertainties on each parameter given the observed SED and known observational effects. To accomplish this goal, we present the probabilistic framework for the Bayesian Extinction and Stellar Tool (BEAST). The development of this methodology has been motivated by the PHAT program (Dalcanton et al. 2012b), an 828-orbit HST multi-cycle program that has observed  $\sim 1/3$  of Andromeda’s star-forming disk from the UV through the near-infrared (NIR). PHAT has cataloged 6-band fluxes for  $> 100$  million individual stars (Williams et al. 2014), forming a critical dataset for better understanding the relationship between star formation and the evolution of M31’s baryonic content. The guiding principle of the BEAST is to accurately derive the stellar and dust parameters in all signal-to-noise regimes, and report well characterized uncertainties. Thus results from the BEAST allow for the study of individually well measured stars as well as statistical studies that take advantage of the large number of sources detected in modern surveys.

We start in §2 with the details of our fitting technique, explicitly including correlations in the uncertainties between the observed bands using a multivariate Gaussian distribution. In §3, we provide the details of our model for stellar SEDs and develop a mixture model for the dust extinction. As an example of applying the BEAST to a specific dataset, in §4 we give the implementation details for the PHAT data. Examples of BEAST results using PHAT data are given in §5.

## 2. FITTING TECHNIQUE

We undertake a probabilistic approach to modeling the SED of a single star. Probability theory provides an established framework for merging the information of multiple separate models (e.g., physical models of stars and dust), comparing the resulting model with observations, and tracking all sources of uncertainty, including covariances between parameters.

Our data consist of  $N$  photometric measurements of a single source. We designate this set of flux measurements as  $F_D$ , and then write the probability of observing  $F_D$  given our model parameters,  $\theta$ , using a multivariate Normal/Gaussian distribution as

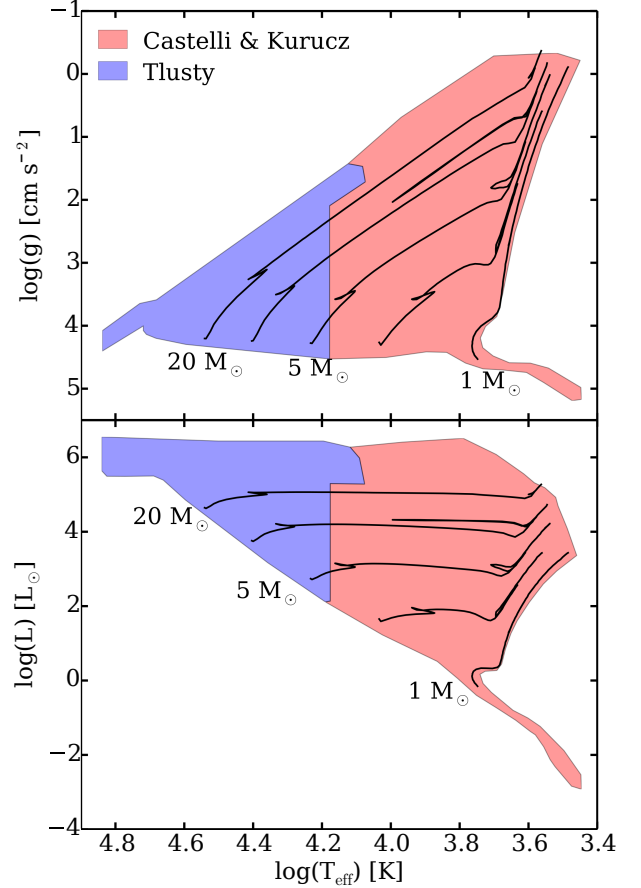
$$P(F_D | \theta) = \frac{1}{Q(\theta)} \exp\left(-\frac{1}{2}\chi^2(\theta)\right), \quad (1)$$

where

$$Q^2(\theta) \equiv (2\pi)^N \det |\mathbb{C}(\theta)|, \quad (2)$$

$$\chi^2(\theta) = \Delta^T \mathbb{C}(\theta)^{-1} \Delta, \quad (3)$$

$$\Delta = F_D - F_M(\theta) + \mu(\theta), \quad (4)$$



**Figure 2.** The coverage of the stellar models used in the BEAST are shown in a Hertzsprung-Russell diagram (top) and  $\log(T_{\text{eff}})$ – $\log(g)$  plot (bottom). The two stellar atmosphere models used are Castelli & Kurucz (2004) and TLusty (Lanz & Hubeny 2003, 2007). Padova stellar evolutionary tracks (Marigo et al. 2008; Girardi et al. 2010) for representative stellar masses are shown as solid black lines for reference.

$\mathbb{C}(\theta)$  is the covariance matrix of the  $N$  photometric bands<sup>15</sup>,  $F_M(\theta)$  is the predicted flux in the  $N$  photometric bands, and  $\mu(\theta)$  is the crowding bias in each band. The use of a multivariate Normal/Gaussian function to compute  $\chi^2$  accounts for the correlations in the observed measurements between bands, unlike the usual assumption that measurements are independent between bands.

The covariance matrix  $\mathbb{C}(\theta)$  is a  $N \times N$  matrix and can be conceptually thought of as a combination of three components such that

$$\mathbb{C}(\theta) = \mathbb{C}_P(\theta) + \mathbb{C}_\mu(\theta) + \mathbb{C}_C(\theta) \quad (5)$$

where  $\mathbb{C}_P(\theta)$  is a diagonal covariance matrix of the photon counting uncertainties<sup>16</sup>,  $\mathbb{C}_\mu(\theta)$  is the covariance matrix due to crowding uncertainties, and  $\mathbb{C}_C(\theta)$  is the covariance matrix giving absolute flux calibration uncertainties. Correlations between bands are significant for

<sup>15</sup> We have chosen to use the notation  $\mathbb{C}(\theta)$  for the covariance matrix instead of the more standard  $\Sigma$  to avoid confusion with the use of  $\Sigma$  as the summation symbol.

<sup>16</sup> In the regime of many photons.



the crowding and absolute flux uncertainties (§4.4). It is worth noting that eq. 3 reduces to the standard  $\chi^2$  equation (e.g., eq. 12.11 in Taylor 1997) for a diagonal covariance matrix (independent measurements).

Having established our likelihood function, we use Bayes's rule to write the probability of the model parameters given the observations, i.e., the posterior probability distribution function (pPDF), as

$$P(\theta | F_D) \propto P(F_D | \theta) P(\theta) \quad (6)$$

where the prior  $P(\theta)$  reflects any external or additional independent information placed on the model parameters. We discuss reasonable priors and illustrate them in the context of the PHAT survey in §4.3.

Our fitting technique assumes  $F_D$  is fully populated with measurements (e.g., there are no upper limits). This avoids the computational complexity of properly accounting for upper limits in the fitting and produces fitting that reflects the full measurements. Modern photometry codes routinely produce flux measurements in all bands for any detected source (e.g., Dolphin 2000; Anderson et al. 2008) removing the burden to account for upper limits in SED fitting.

### 3. DUST EXTINGUISHED STELLAR MODEL

The physical model used in the BEAST provides predictions of the SED of a single star extinguished by dust.

#### 3.1. Single Star Intrinsic SED

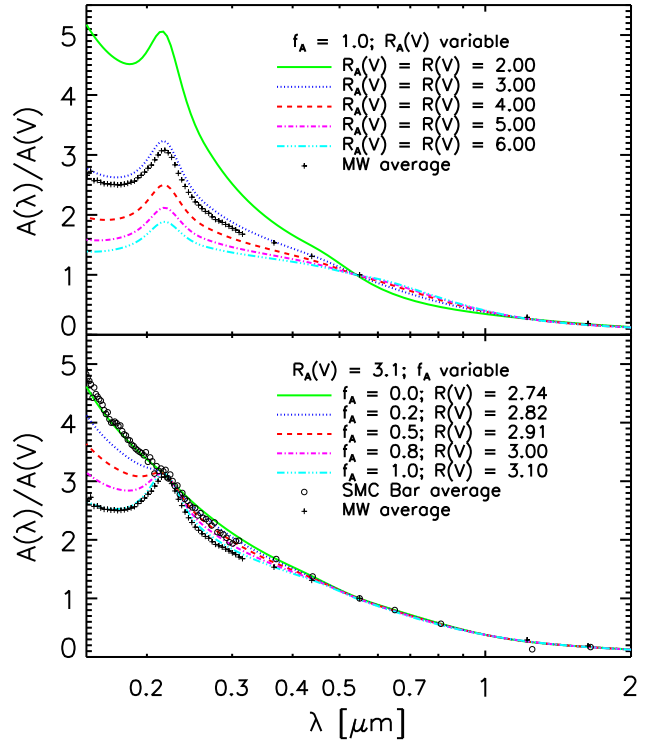
The physics of a star's intrinsic SED can be described by a combination of stellar atmosphere and stellar evolutionary models. The wavelength dependent luminosity, i.e., its SED, of a single star with birth mass  $M$ , age  $t$ , and metallicity  $Z$  is

$$L_\lambda(M, t, Z) = 4\pi [R(M, t, Z)]^2 \times S_\lambda(T_{\text{eff}}, \log(g), Z), \quad (7)$$

where  $R(M, t, Z)$  is the star's radius, given by stellar evolution models, and  $S_\lambda(T_{\text{eff}}, \log(g), Z)$  is the star's surface flux given by stellar atmosphere models and parameterized by the star's effective temperature,  $T_{\text{eff}}$ , surface gravity,  $\log(g)$ , and metallicity  $Z$ . Stellar evolution models provide a mapping of  $(M, t, Z)$  to  $(T_{\text{eff}}, \log(g), Z)$ , effectively reducing the number of free parameters needed to fully describe a star's SED. We have selected  $(M, t, Z)$  as the fundamental stellar parameters because there is a direct mapping given by the stellar evolution models between  $(M, t, Z)$  and  $(T_{\text{eff}}, \log(g), Z)$ , whereas the reverse has degeneracies such that at the same  $(T_{\text{eff}}, \log(g))$  can have multiple possible values of  $(M, t)$ . Thus, we write the fundamental stellar parameters as:

$$\theta_{\text{star}} = \{M, t, Z\}. \quad (8)$$

The BEAST uses a merger of two popular stellar atmosphere grids, the local thermal equilibrium (LTE) CK04 grid (Castelli & Kurucz 2004) and non-LTE TLusty OS-TAR and BSTAR grids (Lanz & Hubeny 2003, 2007). The merging was done by preferring the TLusty non-LTE models over the CK04 LTE models in regions of overlap, given the higher fidelity of the non-LTE modeling for hot stars although the spectra of stars in the overlap regions are very similar (Lanz & Hubeny 2003). We confirmed



**Figure 3.** The behavior of the new, expanded two parameter ( $R(V)$ ,  $f_A$ ) model for the normalized dust extinction curves is illustrated. The top panel gives the variation in the curves as a function of  $R_A(V)$  for the fixed value of  $f_A = 1$ . This shows the  $A$  component of the model. Overplotted is the average extinction curve for all the Milky Way sightlines studied by Gordon et al. (2009). The bottom panel gives the variation in the curves as a function of  $f_A$  for a fixed value of  $R_A(V) = 3.1$ . Overplotted is the same Milky Way average curve as in the top panel as well as the SMC Bar average extinction curve (Gordon et al. 2003) that is the  $B$  component of our model.

that both grids give very similar spectra for the same atmospheric parameters in the overlap region as has been noted previously (Lanz & Hubeny 2003). This provides a seamless merged grid that has excellent  $(T_{\text{eff}}, \log(g))$  coverage, as illustrated by Fig. 2. The stellar evolutionary tracks used by the BEAST are the Padova (Marigo et al. 2008; Girardi et al. 2010) or PARSEC (Bressan et al. 2012, 2013) tracks used in the form of isochrones downloaded from the CMD website with no additional interpolation.<sup>17</sup> The coverage of the stellar atmosphere and evolutionary track models is visualized two different ways in Fig. 2. Potential future expansions of the BEAST stellar model would be the addition of stellar atmosphere models to cover a wider range of stars (e.g., Westera et al. 2002; Bergemann et al. 2012; Rauch et al. 2013) and using newer stellar evolutionary tracks as they become available (e.g., Chen et al. 2014; Tang et al. 2014).

#### 3.2. Interstellar Dust Extinction

Interstellar dust extinguishes stellar light as it travels from the star's surface to the observer. The wavelength-dependence of the extinction from the UV to the NIR has been measured along many sightlines in the Milky Way (Cardelli et al. 1989; Fitzpatrick 1999; Valencic et al. 2004; Gordon et al. 2009) and for a handful of sightlines

<sup>17</sup> <http://stev.oapd.inaf.it/cgi-bin/cmd>

in the Magellanic Clouds (Gordon & Clayton 1998; Missett et al. 1999; Maíz Apellániz & Rubio 2012) as well as in M31 (Bianchi et al. 1996, Clayton et al. 2015, submitted). The observations show a wide range of dust column normalized extinction curves,  $A(\lambda)/A(V)$ . Here we introduce a mixture model with two components  $\mathcal{A}$  and  $\mathcal{B}$  to describe the full range of observed extinction curves in the Local Group.

For most Milky Way sightlines, the variations in dust extinction at a particular  $\lambda$  have been found to be, on average, linearly dependent on the single parameter  $R(V) = A(V)/E(B-V)$  (Cardelli et al. 1989; Fitzpatrick 1999; Gordon et al. 2009). We adopt this Milky Way extinction model as the  $\mathcal{A}$  component of our mixture model, expressed as

$$\left[ \frac{A(\lambda)}{A(V)} \right]_{\mathcal{A}} = a(\lambda) + \frac{b(\lambda)}{R_{\mathcal{A}}(V)}, \quad (9)$$

where the equations for  $a(\lambda)$  and  $b(\lambda)$  are determined from linear fits to  $A(\lambda)/A(V)$  versus  $R(V)^{-1}$  (Cardelli et al. 1989). We use Fitzpatrick (1999) to compute  $[A(\lambda)/A(V)]_{\mathcal{A}}$  as it is explicitly formulated to account for passband effects of the optical and NIR extinction curve measurements. The behavior of the  $\mathcal{A}$  component is shown in the top panel of Fig. 3.

Most sight-lines in the Magellanic Clouds do not follow the Milky Way  $R(V)$ -dependent relationship (Gordon et al. 2003). For example, the extinction curves in the Small Magellanic Cloud (SMC) star-forming Bar lack the usually strong 2175 Å extinction bump and show a wavelength dependence that is nearly linear versus  $\lambda^{-1}$ . We have found that the Magellanic Cloud and “deviant” Milky Way sightlines (Mathis & Cardelli 1992; Valencic et al. 2004) can be represented by a mixture model given by

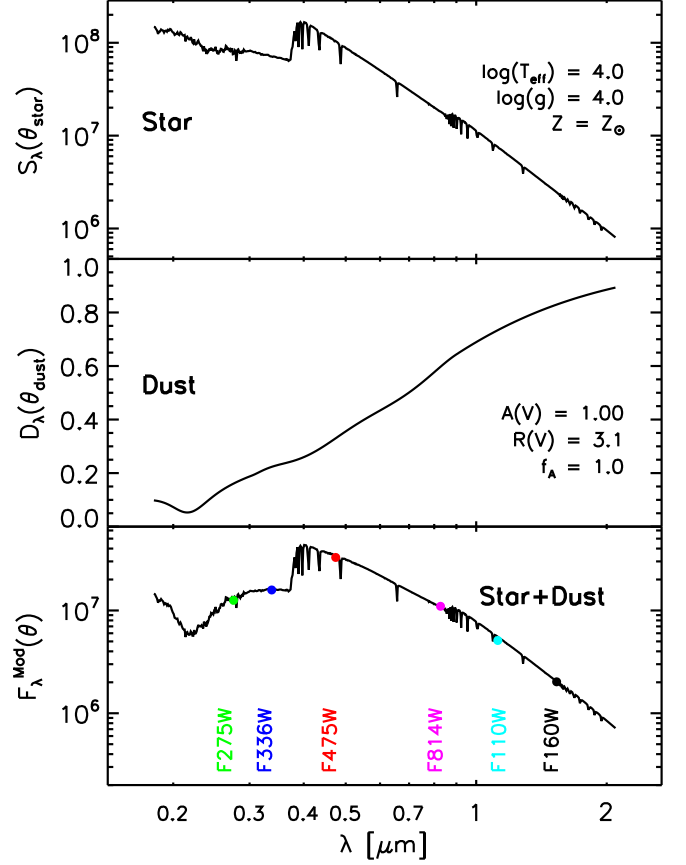
$$k_{\lambda}(R(V), f_{\mathcal{A}}) = f_{\mathcal{A}} \left[ \frac{A(\lambda)}{A(V)} \right]_{\mathcal{A}} + (1 - f_{\mathcal{A}}) \left[ \frac{A(\lambda)}{A(V)} \right]_{\mathcal{B}}, \quad (10)$$

where  $f_{\mathcal{A}}$  gives the fraction of the  $\mathcal{A}$ -type extinction and  $(1 - f_{\mathcal{A}})$  the fraction of the  $\mathcal{B}$ -type extinction (Tchernyshyov & Gordon, in prep.). The wavelength dependence of the  $\mathcal{B}$ -type extinction is given using the SMC Bar average UV parameters and optical/NIR data points from Gordon et al. (2003) using the technique of Fitzpatrick (1999) to smoothly interpolate between band extinctions in the optical and NIR. The  $J$  and  $K$  band values of  $A(\lambda)/A(V)$  were changed from those given by Gordon et al. (2003) to 0.25 and 0.11, respectively, to provide a smooth, non-negative cubic spline interpolation. Note that  $f_{\mathcal{A}}$  is not a direct measure of the 2175 Å extinction feature, but is a measure of one component of the extinction curve shape variation across the full wavelength range. The behavior of this mixture model is illustrated in Fig. 3.

The  $R(V)$  of the mixture extinction curve model is

$$R(V)^{-1} = f_{\mathcal{A}} R_{\mathcal{A}}(V)^{-1} + (1 - f_{\mathcal{A}}) R_{\mathcal{B}}(V)^{-1} \quad (11)$$

where we fix  $R_{\mathcal{B}}(V) = 2.74$  (Gordon et al. 2003). The range of observed  $R_{\mathcal{A}}(V)$  is between 2.0 and 6.0 and this results in the parameter space defined by  $(R(V), f_{\mathcal{A}})$  not being completely filled, see §4.3.



**Figure 4.** The computation of a BEAST model SED for a dust extinguished star is shown graphically. The top panel gives the stellar spectrum, the next panel shows the extinction by dust, and the bottom panel plots the full extinguished stellar spectrum. In the bottom panel, the integrated SEDs for the PHAT bandpasses are plotted at their effective wavelengths ( $\lambda_{\text{eff}}$ , solid circles).

With this dust mixture model, interstellar extinction is included in our model with the multiplicative term

$$D_{\lambda, \text{dust}}(\theta_{\text{dust}}) = 10^{-0.4 A(V) k_{\lambda}(R(V), f_{\mathcal{A}})} \quad (12)$$

that has the three dust parameters

$$\theta_{\text{dust}} = \{A(V), R(V), f_{\mathcal{A}}\} \quad (13)$$

where  $A(V)$  is the extinction in magnitudes in the Johnson  $V$  band and the  $(R(V), f_{\mathcal{A}})$  parameter combination defines the shape of the extinction curve.

### 3.3. Full SED Model

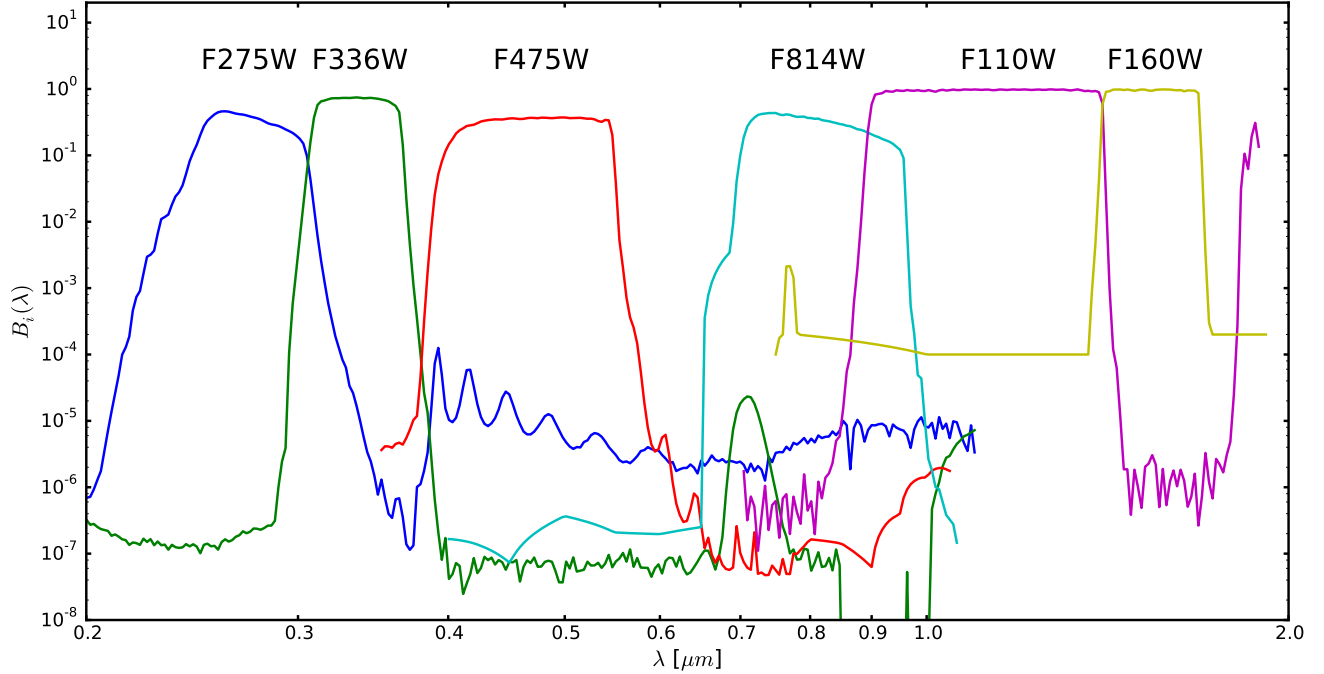
The results of the previous two subsections (Eq. 7 and 12) provide a model of a star’s observed monochromatic flux after passing through a column of dust. For a star at a distance  $d$ , we write the full model for the monochromatic flux  $F_{\lambda}^{\text{Mod}}$  as

$$F_{\lambda}^{\text{Mod}}(\theta) = \frac{L_{\lambda}(\theta_{\text{star}}) D_{\lambda, \text{dust}}(\theta_{\text{dust}})}{4\pi d^2} \quad (14)$$

where

$$\theta = \{\theta_{\text{star}}, \theta_{\text{dust}}, d\} \quad (15)$$

$$= \{M, t, Z, A(V), R(V), f_{\mathcal{A}}, d\}. \quad (16)$$



**Figure 5.** The band response functions for the six PHAT filters are plotted. The plot is shown with a log scale to illustrate the response seen over the full wavelength range (i.e. showing the red and blue “leaks”). These response functions include both the filter throughputs and the detector efficiencies.

To compare with photometric observations, we need to compute model fluxes in the same bands as the observations. We calculate the model band flux in bandpass  $i$  using

$$F_i^{\text{Mod}} = \frac{\int \lambda B_i(\lambda) F_\lambda^{\text{Mod}}(\theta) d\lambda}{\int \lambda B_i(\lambda) d\lambda}, \quad (17)$$

where  $B_i(\lambda)$  is the bandpass response function for the  $i$ th band in fractional photon units. This integration is done in photon units (via  $\lambda d\lambda$ ) to correctly model how the measurements were made (e.g., photon-based detectors Hogg et al. 2002; Sirianni et al. 2005). We explicitly calculate the flux in each band from the the response functions and model spectra including dust extinction, removing the need to deal with color and bolometric corrections not associated with inaccuracies in the stellar models.

Fig. 4 gives a graphical representation of how the model SEDs are computed showing the intrinsic spectrum, the effects of dust, the extinguished spectrum, and the band integrated SED.

#### 4. PHAT IMPLEMENTATION DETAILS

As a concrete example of the use of the BEAST on a large set of resolved stellar photometry, we fit the 0.7 million stars detected in at least 4 bands in Brick 21 of the PHAT survey. Preparing and running the BEAST on this data set requires several steps of implementation, which are described in this section.

##### 4.1. Grid Implementation

We have implemented the BEAST probabilistic fitting using a grid-based approach. This approach ensures that the entire topology of the posterior function is explored (i.e., the full model parameter space  $\theta$ ). For realistic

sampling of the model grid this approach is faster than a Monte Carlo Markov Chain (MCMC) or nested sampling approach. Finally it avoids the computationally intensive normalizations needed for MCMC results for further use of the results (e.g., hierarchical models of stellar clusters, dust geometry, etc.). For the PHAT survey, the final range and grid spacing of each parameter is given in Table 1.

##### 4.2. Response Functions

For PHAT, the observed bands are HST/WFC3 F275W, F336W, F110W, and F160W and HST/ACS F475W and F814W. The full bandpass response functions used in creating the model SEDs are shown in Fig. 5. The bandpass response functions are for the full instrument plus telescope system including the detector response. This figure shows that these photometric bands can be quite broad. For such broad photometric bands, it is important to include the effects of dust extinction before integrating across the photometric bands, given that the SED spectral shape changes the effective wavelength of the measurement.

Integrating the dust-free stellar SED and then multiplying by the dust extinction using a constant  $A(\text{band})/A(V)$  would not only be formally incorrect, but it also leads to large errors in the model band fluxes. For example, applying the dust extinction after band integration for BEAST models bright enough to be formally detected in the PHAT survey with  $R(V) = 3.0$  and  $f_A = 1.0$  would result in errors  $> 10\%$  for 52%, 1%, 24%, 25%, 24%, and 0% for the F275W, F336W, F475W, F814W, F110W, and F160W bands, respectively. The maximum error for the same models is 56%, 15%, 45%, 34%, 29%, and 3% for the same bands. These results are determined by three factors: the largest  $A(V)$  de-

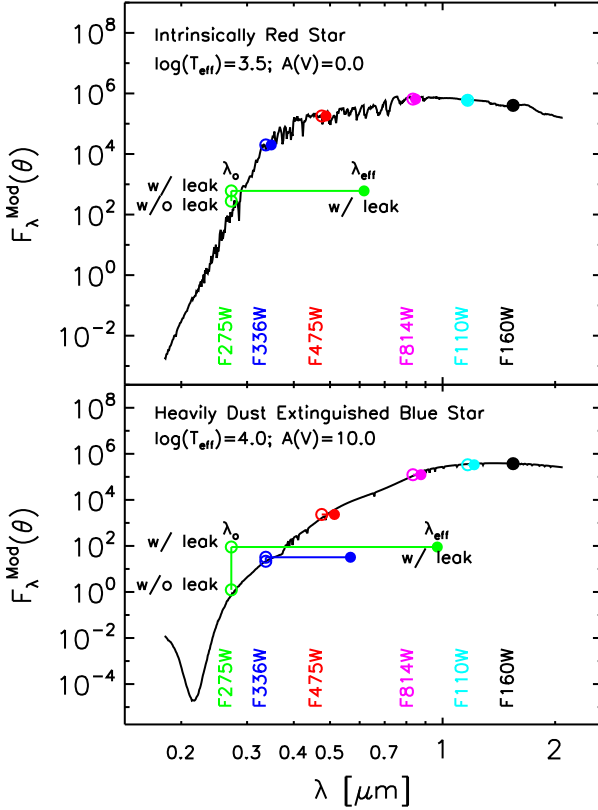
**Table 1**  
Model Parameters

Parameter	Description	Min	Max	Resolution	Prior
$\log(t)$ [years]	stellar age	6.0	10.13	0.05	flat SFR
$\log(M)$ [ $M_{\odot}$ ]	stellar mass	-0.8	2.0	variable <sup>a</sup>	Kroupa IMF <sup>b</sup>
$\log(Z)$ <sup>c</sup>	stellar metallicity	-2.3	0.1	0.1	flat
$A(V)$ [mag]	dust column	0.0	10.0	0.02	flat
$R(V)$	dust average grain size	2.0	6.0	0.5	peaked at $\sim 3$ (Fig. 7)
$f_A$	dust mixture coefficient	0.0	1.0	0.1	peaked at 1 (Fig. 7)
$d$	distance		776 kpc		$\delta$ function

<sup>a</sup> Determined by stellar lifetime and fair sampling of stellar evolutionary phases.

<sup>b</sup> In future work, we will use the Weisz et al. (2015) updated M31 IMF.

<sup>c</sup>  $Z$  is the mass fraction of all elements other than hydrogen and helium.



**Figure 6.** The spectra and photometric SEDs for an intrinsically red star and a heavily dust extinguished blue star are plotted as computed by the BEAST. The band integrated SEDs are plotted at their effective wavelengths ( $\lambda_{\text{eff}}$ , solid circles) and the nominal wavelengths ( $\lambda_o$ , open circles). To illustrate the impact of the red leaks, additional open circles labeled with “w/o leak” for the F275W and F336W bands are plotted giving the fluxes computed where the red leaks are removed (i.e., zeroing the band response functions at  $\lambda > 0.37$  and  $0.4 \mu\text{m}$ , respectively). The model parameters not indicated on the plots are  $\log(g) = 4.0$ ,  $Z_{\odot}$ ,  $R(V) = 3.1$ , and  $f_A = 1$ .

tectable at the survey depth in a band increases with wavelength; the increasing  $A(\text{band})/A(V)$  with decreasing wavelength; and the width of a band’s response function. Overall, the trend is for the errors to decrease as the nominal wavelength of the band increases. The notable exceptions to these trends are for F336W and F160W which have the lowest errors as these bands having significantly narrower band response functions than the other four bands (see Fig. 5).

In addition to the broad band response functions, some

of the bands also have red or blue leaks (e.g., F336W with a red leak at  $\sim 0.7 \mu\text{m}$ , F110W with a strong red leak at  $\sim 1.9 \mu\text{m}$ , and F160W with a blue leak at  $\sim 0.8 \mu\text{m}$ ). The impact of such leaks is illustrated in Fig. 6 for an intrinsically red star and a heavily dust extinguished blue star. The band integrated fluxes for a number of the filters are influenced by the steep spectrum inside the main bandpass and the two bluest filters are strongly influenced by the contribution from the filter red leak. The impact of red leaks are strong for both stars’ F275W fluxes and for the F336W flux of the reddened blue star. These predictions illustrate that for such red sources, the bluest filters will collect more photons than nominally expected due to the red leak signal being much larger than the main band signal.

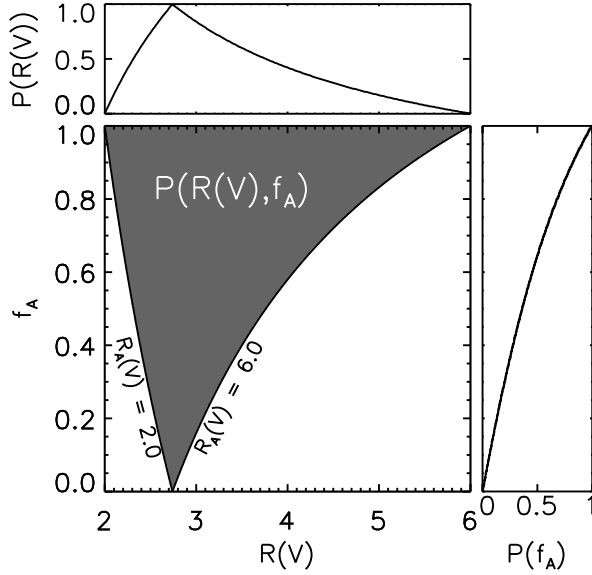
One way to quantify the impact of the leaks is to calculate the effective wavelength of the predicted fluxes. The effective wavelength ( $\lambda_{\text{eff}}$ ) is the flux throughput weighted average wavelength, while the nominal wavelength ( $\lambda_o$ ) is the throughput only weighted average wavelength. The effective wavelengths show that for the intrinsically red star, the flux measured in what is a UV filter (F275W) is actually due to optical photons. For the reddened blue star, the effect is even larger with the F275W flux dominated by NIR photons. The filter fluxes for sources strongly impacted by leaks are often found below the actual SED at  $\lambda_{\text{eff}}$ . In these cases, the measured flux can be thought of as an average of flux in the main band and the leak band that has a much lower throughput than the main band with an effective wavelength somewhere in between the main and leak bands.

While the examples given in Fig. 6 were chosen to be extreme to illustrate the impact of red leaks, a portion of the parameter space will be impacted by such leaks at a measurable level. For the models that have F275W fluxes detectable with the PHAT observations (i.e.,  $> 5 \times 10^{-19} \text{ ergs cm}^{-2} \text{ s}^{-1} \text{ \AA}^{-1}$ ), the F275W red leak has a larger than 5% effect for 9% of the BEAST models with a maximum impact of  $\sim 50\%$ . For the models that have F336W fluxes detectable with the PHAT observations (i.e.,  $> 1 \times 10^{-19} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ \AA}^{-1}$ ), the F336W red leak has a larger than 5% effect for 0.6% of the BEAST models with a maximum impact of  $\sim 14\%$ .

Thus, correctly performing the filter integrations after fully generating the model of a dust extinguished star is critical to achieve high precision and avoid systematic biases in the resulting model fit parameters.

#### 4.3. Priors



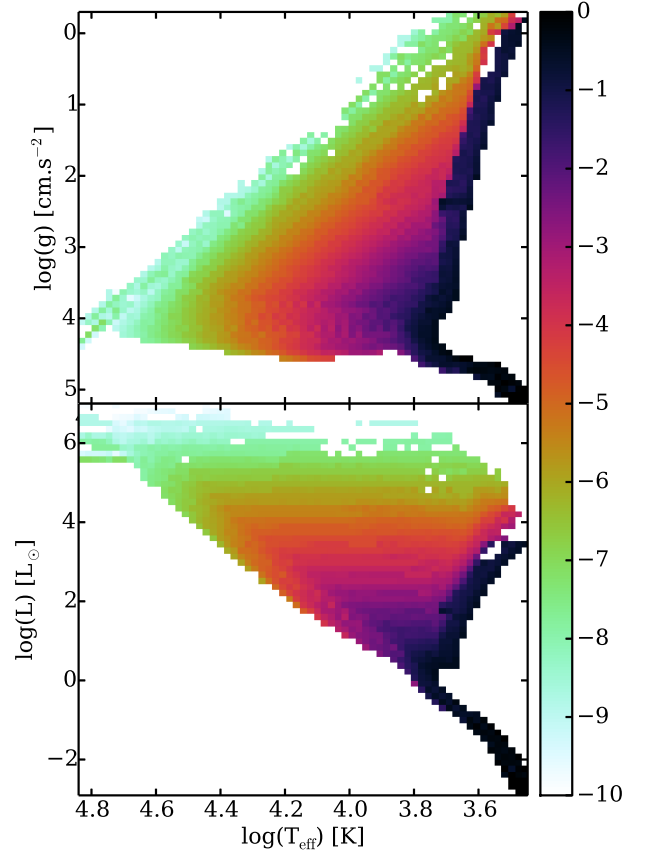


**Figure 7.** The prior on  $R(V)$  and  $f_A$  is shown, both in 2D and in 1D forms. The prior is set by imposing a fixed range in  $R_A(V)$  from 2 to 6 as this corresponds to the observed range for the Milky Way sightlines used to define the  $\mathcal{A}$  component (Cardelli et al. 1989; Fitzpatrick 1999; Gordon et al. 2009).

All fitting includes priors, whether they are acknowledged or not. Explicitly using priors on the model parameters provides a clear way to acknowledge and quantify what are often seen as assumptions in the fitting process. Priors also provide a quantitative way to incorporate independent knowledge about the stars and dust in a galaxy from previous studies in the fitting. For example, imposing a prior on the stellar masses by using an Initial Mass Function (IMF) can greatly help the statistical accuracy of the fitting, since the known steep IMF favors the production of low mass stars (see the review of Bastian et al. 2010). Priors have a significant effect when the data does not provide strong constraints on a parameter. For example, priors have a strong impact on low signal-to-noise measurements but only a minor impact on high signal-to-noise measurements. Given that the BEAST has been developed to fit survey observations where the majority of detected sources have a low signal-to-noise, it is important to carefully choose the priors.

The BEAST fitting and marginalization is done using a grid due to the complexity of the physical model and need for computational speed. A uniform grid in all dimensions provides for the easiest and fastest implementation as the integration needed to marginalize over any parameter becomes a simple summation as all points have the same nD volume. Computational and memory considerations motivate a logarithmic or non-uniform spacing for some of the parameters. We account for the such spacing and impose reasonable and physically motivated priors using weights.

The priors that result from the final weighting are summarized in Table 1. We adopt a flat prior on  $A(V)$ , as we have no prior information about the expected  $A(V)$  distribution. We implement this prior using equal spacing of the  $A(V)$  points in the model grid. We adopt a uniform prior in the 2D space of  $R(V)$  versus  $f_A$  subject to the constraint that  $2 < R_A(V) < 6$  as shown in Fig. 7.



**Figure 8.** The mapping of the stellar  $t$  (age) and  $M$  (mass) priors into the  $\log(T_{\text{eff}})$  versus  $\log(L)$  and  $\log(T_{\text{eff}})$  versus  $\log(g)$  spaces are shown in a log scaling with the peak weight set to 1.0. We use cubehelix color scalings for this figure and throughout our paper as such color scales are robust for various types of color blindness (Green 2011).

This produces a 1D projection of the prior weighted towards high values of  $f_A$  and  $R(V) \sim 3$ . The resulting  $R(V)$  and  $f_A$  priors are reasonable as they roughly reflect the range of observed  $R(V)$  values (e.g., Fig. 3 in Gordon et al. 2009) and the strong dominance of  $f_A \sim 1$  curves in the total sample of observed extinction curves (Gordon et al. 2003; Valencic et al. 2004).

The intrinsic grid sampling in stellar mass is driven by the need for the stellar evolutionary model outputs to efficiently sample the evolutionary phases. This sampling is quite different from that of a reasonable IMF. We use a multiplicative weight on the grid points to achieve a Kroupa IMF (Kroupa 2001) prior at all ages. We add an additional multiplicative weight that is a function of age ( $t$ ) to give a uniform prior, as the intrinsic  $t$  grid points are logarithmically spaced for computational speed. Finally, we add a third multiplicative term to impose a flat prior on the stellar metallicity ( $Z$ ) distribution. The mapping of the  $M$  and  $t$  priors into Hertzsprung-Russell and stellar atmosphere  $\log(T_{\text{eff}})$  versus  $\log(g)$  diagrams is shown in Fig. 8. The final priors reflect the expected distribution of real stars, with high densities of low mass stars and older red giant branch (RGB) stars.

The final prior is on M31's distance for which we as-



sume a value of 776 kpc (as adopted by Dalcanton et al. 2012c). This choice is well justified given that the distance to M31 is well measured and the high density of M31 stars ensures very small contamination by MW foreground sources. The potential variation in the distance to M31 due to the depth of the galaxy is also small (e.g., a 20 kpc radius is  $<3\%$  of M31’s distance). We tested the sensitivity of our results to the assumed distance. We find that the recovered parameters (within the  $1\sigma$  confidence intervals) do not change with distance variations on the order of 10% for the PHAT bands and survey depth.

#### 4.4. Noise Model

The noise model in the BEAST is defined by the bias vector  $\mu$  and the covariance matrix  $\mathbb{C}$  which is composed of three terms: photon, crowding, and absolute flux calibration (Eq. 5).

##### 4.4.1. Absolute Flux Term

The wavelength dependent absolute calibration of HST is based on the average of spectroscopic measurements of the predicted to observed ratios of three white dwarf stars. The full details are given by Bohlin et al. (2014) and the details relevant to this work are summarized here. The predicted spectrum for each white dwarf is based on ground-based spectroscopic measurements of their stellar atmosphere parameters (e.g.,  $T_{\text{eff}}$ ,  $\log(g)$ , etc.). This predicted spectrum is scaled by the measured ratio of their fluxes to that of Vega at 5556 Å and the absolute measurement of Vega’s flux at this wavelength compared to laboratory calibrated blackbodies. The absolute flux calibration covariance matrix  $\mathbb{C}_C$  for HST has been derived based on uncertainties in the calibration steps (Bohlin et al. 2014). The HST  $\mathbb{C}_C$  is composed of two components. The first captures the uncertainties and covariance associated with variations in the spectral shape of the stellar atmosphere models of the three white dwarfs (i.e., uncertainties in  $\log(T_{\text{eff}})$ ,  $\log(g)$ , etc.). The second component is a uniform 0.7% fully correlated uncertainty due to the uncertainty in measurements of the absolute flux scale via Vega’s flux at 5556 Å. Following the recommendations by Bohlin et al. (2014), we generated the photometric band covariance matrix for each model by averaging the spectroscopic resolution matrix using the PHAT filter response functions and adding in quadrature the 0.7% uncertainty to all matrix elements. The terms of the final absolute flux covariance matrix are

$$\mathbb{C}_C^{ij} = F_M^i(\theta) F_M^j(\theta) \mathbf{A}^{ij}(\theta) \quad (18)$$

where  $F_M(\theta)$  is the predicted flux in the  $N$  photometric bands (eq.17) and the  $\mathbf{A}(\theta)$  matrix consists of the fractional uncertainties. Fractional values are the natural units of the Bohlin et al. (2014) results given the multiplicative nature of the absolute calibration. We illustrate the range in the  $\mathbf{A}(\theta)$  matrix over the model grid by giving the minimum and maximum values for each term in

the matrix. Thus,

$$100\mathbf{A}(\theta)_{\min}^{1/2} = \begin{bmatrix} 0.70 & 0.50 & 0.69 & 0.51 & 0.30 & 0.23 \\ 0.50 & 0.70 & 0.69 & 0.54 & 0.40 & 0.37 \\ 0.69 & 0.69 & 0.70 & 0.69 & 0.67 & 0.67 \\ 0.51 & 0.54 & 0.69 & 0.72 & 0.74 & 0.74 \\ 0.30 & 0.40 & 0.67 & 0.74 & 0.83 & 0.85 \\ 0.23 & 0.37 & 0.67 & 0.74 & 0.85 & 0.88 \end{bmatrix} \quad (19)$$

and

$$100\mathbf{A}(\theta)_{\max}^{1/2} = \begin{bmatrix} 1.20 & 1.11 & 0.75 & 0.79 & 0.83 & 0.85 \\ 1.11 & 1.05 & 0.74 & 0.79 & 0.83 & 0.85 \\ 0.75 & 0.74 & 0.71 & 0.70 & 0.70 & 0.70 \\ 0.79 & 0.79 & 0.70 & 0.76 & 0.79 & 0.80 \\ 0.83 & 0.83 & 0.70 & 0.79 & 0.84 & 0.86 \\ 0.85 & 0.85 & 0.70 & 0.80 & 0.86 & 0.89 \end{bmatrix}, \quad (20)$$

where we have multiplied  $\mathbf{A}$  by 100 for display purposes. The diagonal terms show that the overall level of absolute flux uncertainty is around 1%. The non-diagonal terms show that this uncertainty is highly correlated between bands. The minimum and maximum values show variations around a factor of two at shorter wavelengths and much smaller variations in the longer wavelengths. The difference in variations between shorter and longer wavelengths is directly related to larger variations in spectral shape between these wavelength regions.

##### 4.4.2. Combined Photon and Crowding Term

The photon and the crowding terms are combined and measured simultaneously through the use of artificial star tests (ASTs). These tests accurately capture the non-linear interaction between the photon and crowding noise. The crowding not only impacts the noise, but also can systematically bias the flux of a source. Such systematics become dominant in the case of high stellar crowding, when a star’s flux measurement is contaminated by a non-negligible amount of flux from neighboring sources.

ASTs are performed by inserting artificial stars with known SEDs into the observed images and re-running the full photometric pipeline on the altered images. In this way, the fluxes of the artificial stars are extracted using the same technique that is used to produce the observed point source catalog and can be compared to the true input fluxes.

We use the model grid SEDs as inputs for the ASTs. The photon noise is simulated using the parameters of the observations (sensitivity, exposure time, etc.) and each star is inserted into the observed images using the known point spread function. Finally, each star is either measured or undetected as if it were a real source. This process is repeated many times to build up secure statistical information on how photometric measurements are affected by crowding, photon noise, brightness, etc.

The bias  $\mu_i(\theta)$  in band  $i$  given  $N$  ASTs recovered fluxes  $F_R^k$  with an input model  $\theta$  that has intrinsic flux  $F_M(\theta)$  is

$$\mu_i(\theta) = \frac{1}{N} \sum_k (F_{R,i}^k - F_M(\theta)), \quad (21)$$

and the covariance between bands  $i$  and  $j$  is

$$\mathbb{C}_{ij}^2(\theta) = \frac{\sum_k^N (\mu_i^k - \langle \mu_i(\theta) \rangle) (\mu_j^k - \langle \mu_j(\theta) \rangle)}{N - 1}. \quad (22)$$

The bias captures the mean offset between the true and measured fluxes. The covariance captures the magnitude and shape of the scatter about this mean offset. If the different bands are independent of each other, as is often assumed, the entries along the diagonal of the covariance matrix are the squared standard deviations in each band and all off-diagonal entries are equal to zero.

In Fig. 9 we show the AST data and the resulting covariance matrix and bias for a single model SED. This model SED represents a typical PHAT source that is well detected in the redder bands and basically undetected in the three shorter wavelength bands. The highest covariances seen are for F475W, F814W, F110W, and F160W bands. The obvious source of the covariance is neighboring sources (i.e., crowding noise). The impact of crowding noise can be seen even in the F475W band that is detected, on average, only at  $<1\sigma$ , yet the F475W and F814W bands are still strongly correlated. In contrast, the lack of a strong correlation between the two shortest wavelength bands (F275W and F336W) and the redder bands is an indication that the measurements in these bluer bands are dominated by measurement (photon) noise.

Fig. 9 also illustrates that our use of a multi-variate Gaussian to characterize the offset and scatter in the recovered fluxes is an approximation to the true distributions. This can be seen by the asymmetry in the  $\mu_k$  values in many of the bands with more positive deviations than negative deviations. In the future, we will investigate the use of multi-variate skew Gaussians or numerically sampling the AST distribution as a refinement of our noise model. Practical concerns on the computation cost of including enough ASTs per model SED to define the deviations from a multi-variate Gaussian may dominate the discussion. In practice, we expect that the gains in including the asymmetries in the  $\mu_k$  values via a more sophisticated noise model will be smaller than the impact of including covariance to the first order.

The AST derived covariance matrices and biases are dependent on the model SED and on the location of a star in the PHAT survey area. Ideally many ASTs would be run for every BEAST model SED for every pixel (or subpixel position) over the entire survey region. As this is not computationally feasible, we are forced to average the AST results over spatial regions and interpolated between models. The effects of crowding are most strongly dependent on the source density and, thus, we average the AST results over regions of similar source density.

We illustrate the dependence of the AST derived covariance matrices in Fig. 10 for one field in one PHAT Brick. This figure gives projections of the ensemble of the covariance matrices. The uncertainty per band is strongly dependent on the source flux as expected. The correlation between bands is also strongly dependent on flux; with fainter fluxes showing stronger correlations, especially in the longer wavelength bands where crowding is more significant. The step function drop in correlation at the highest fluxes is traced to the fluxes for the brightest stars coming from the the short “guard” exposures, rather than the deeper main survey exposures where the stars are saturated.

#### 4.4.3. Importance of Including Covariance

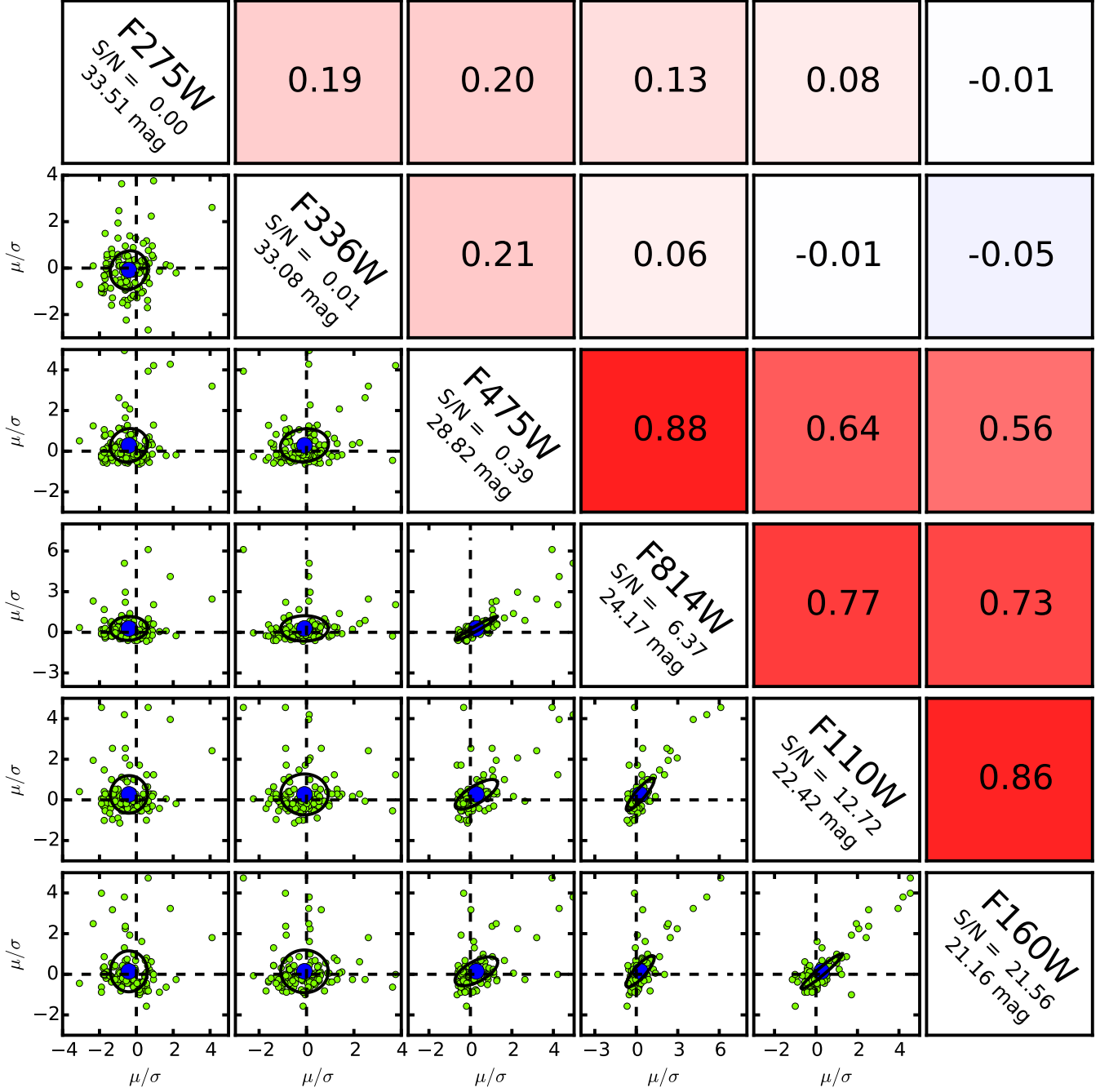
The importance and impact of including the covariance in the SED fitting is illustrated in Fig. 11. This figure shows the 1D pPDFs of the SED fitting likelihood function for a simulated star in Field 15 of PHAT Brick 21. The simulated source is a hot, young star that has experienced appreciable dust extinction. The observational noise was simulated using the full noise model with a covariance matrix interpolated from those measured using a small set of full 6-band ASTs spanning the observed flux range run for this field and brick. The SED fitting was done twice, the first time without using the covariance information (i.e., diagonals only) and the second time using the full covariance matrix. It is clear that including the full noise model with covariance produces a more accurate and precise recovery of true model parameters. Overall, the 1D pPDFs with covariance better recover the input model parameters, given that their peaks are better matched to the input values and their widths are narrower. The differences can be dramatic like those for the  $\log(M)$ ,  $\log(T_{\text{eff}})$ , and  $\log(g)$  model parameters where the 1D pPDFs change from being double peaked without covariance to being dominated by a single peak with covariance. In essence, the covariance restricts the allowed parameter space producing narrow pPDFs.

#### 4.5. Speed Optimizations

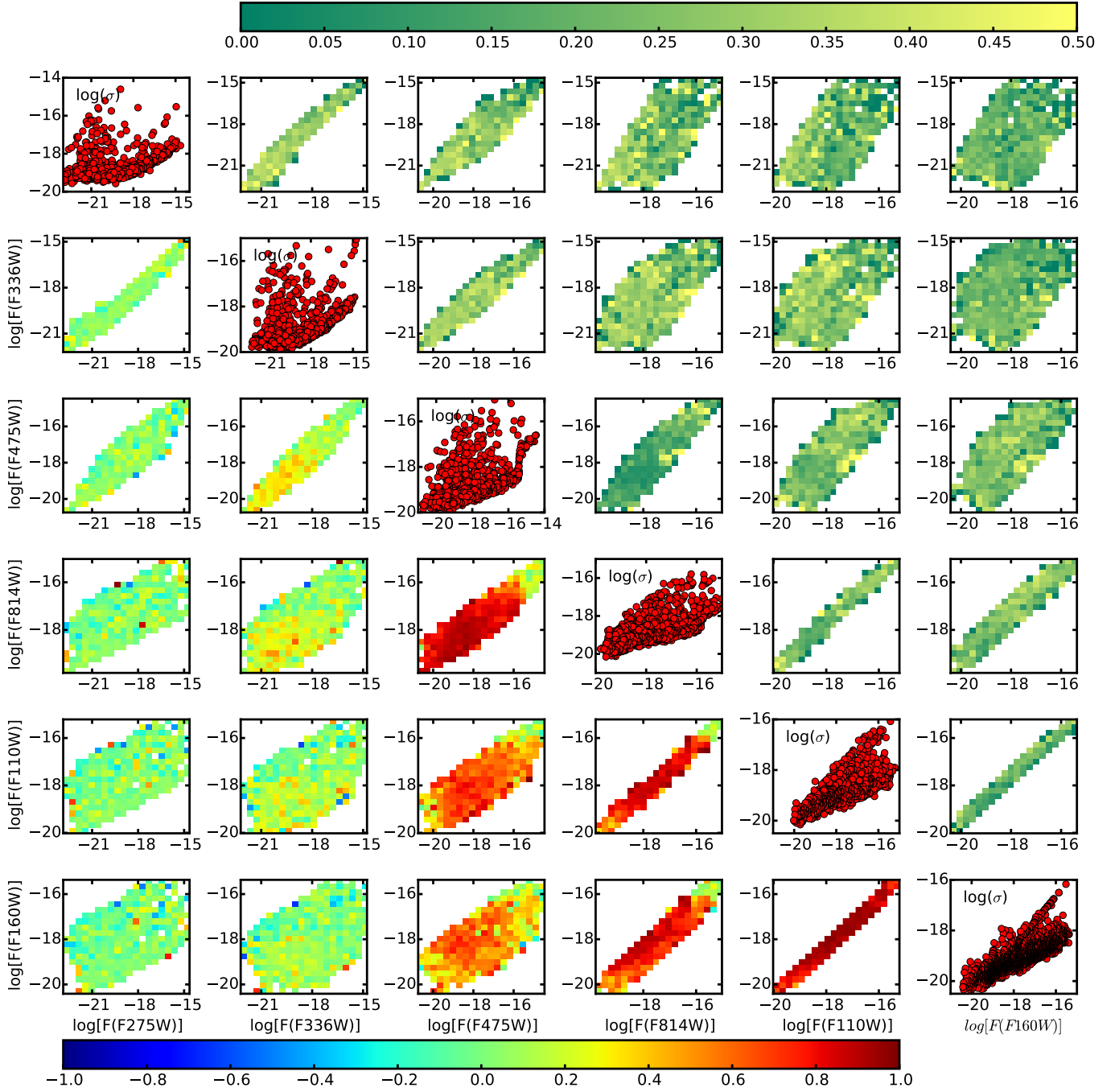
It is fairly quick using modern computers to calculate the full  $N$ -dimensional pPDF for a single observed SED using a grid consisting of millions of models. The challenge is doing this for the  $>100$  million sources in the PHAT survey (Williams et al. 2014). We are pursuing a number of options for speeding up the pPDF calculations, mainly focused on reducing the effective size of the model grid. One straightforward optimization that we have implemented is to only compute the pPDF for models that are in the range of fluxes expected for the survey. Thus, the grid is trimmed of models that predict fluxes that would saturate in the survey observations or are well below the survey sensitivity as measured from the AST results (i.e., with zero completeness). The next potential optimization will be to compress/hash the grid (similar to a tree-code approach) so that the BEAST only computes the pPDF for models with equivalent SEDs within some tolerance (Fouesneau et al., in prep.). We expect this will allow for faster computation rates and/or larger model grids.

### 5. EXAMPLE PHAT RESULTS

To illustrate the BEAST capabilities on real data, we have fit the 0.7 million sources that were detected in at least 4 PHAT bands with a F475W Vega magnitude brighter than 27.6 mag in the PHAT Brick 21 region (Williams et al. 2014). Brick 21 samples a range of star formation regions and dust contents. For this initial work, we use a coarser grid than given in Table 1, specifically with resolutions of 0.15 in  $\log(t)$ , 0.15 in  $A(V)$ , 0.5 in  $R(V)$  and 0.5 in  $f_A$  and  $Z = 0.03, 0.019, 0.008$ , & 0.004. This coarser grid allowed the fitting to be done for the 0.7 million Brick 21 sources in a reasonable amount of time. With this model grid, we find that it takes the BEAST  $\sim 13$  s to fit and save the results for a single star using the Texas Advanced Computing Center (TACC) Stampede supercomputer with time provided through XSEDE (Townes et al. 2014).

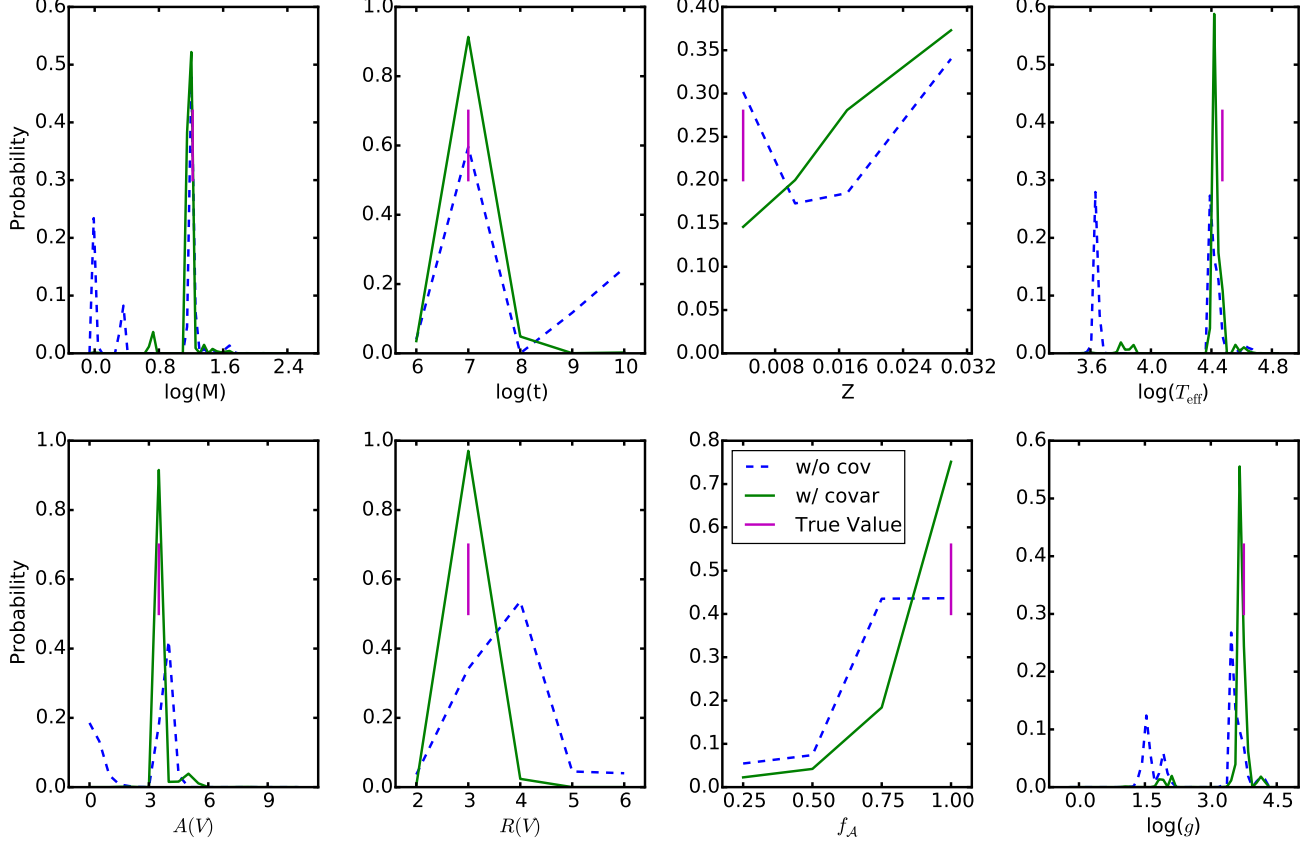


**Figure 9.** The covariance matrix calculation and result is graphically illustrated for a single model SED using 125 ASTs with independent realizations of the photon noise and spatial location in a small region of the PHAT observations. The lower triangle of plots shows the individual biases  $\mu_k$  for each of the  $k$  ASTs as small green circles and the average bias as a larger blue circle. The axes of these plots are the  $\mu_k$  divided by the appropriate band  $\sigma_k$ . The dashed lines give the location of zero bias  $\mu_k$  in each band. The circle is an illustration of the covariance calculated from the individual  $\mu_k$  values. The upper triangle gives the correlation coefficient ( $c_{ij} = \mathbb{C}_{ij}^2 / \sigma_i \sigma_j$ ) for each pair of bands.



**Figure 10.** A visualization of the covariance matrices is shown for Field 2 of PHAT Brick 9. The plots on the diagonal give the  $\sigma_k$  values for each band (square root of the diagonals of the covariance matrix). The plots in the lower triangle (color scale from lower color bar) give the average correlation versus flux in each of the two considered bands. The correlation values are shown instead of covariance values to provide a more understandable interpretation of the values. The plots in the upper triangle (color scale from upper color bar) give the standard deviation in the correlation for each of the two considered bands in each combined flux bin. These plots provide a measure of the additional variation that is due to the fluxes in other bands than the two considered bands.





**Figure 11.** The 1D pPDFs are shown from fitting a simulated source as if it were observed in Field 15 of Brick 21 of the PHAT survey. The star was simulated using the full noise model including covariance. Two sets of 1D pPDFs are shown, one a simple noise model without any covariance information and one using the full covariance noise model. The true values of the simulated source are shown.

The noise model was computed using the artificial star tests (ASTs) discussed by Dalcanton et al. (2012b) and (Williams et al. 2014). Using this noise model, the uncertainty ( $\sigma$ ) and bias ( $\mu$ ) terms as a function of source density are shown in Fig. 12 for two filters to illustrate the different behaviors between photon and crowding limited observations. This figure clearly illustrates the dependence of the bias and uncertainty on flux and source density and how it varies between bands in the PHAT survey.

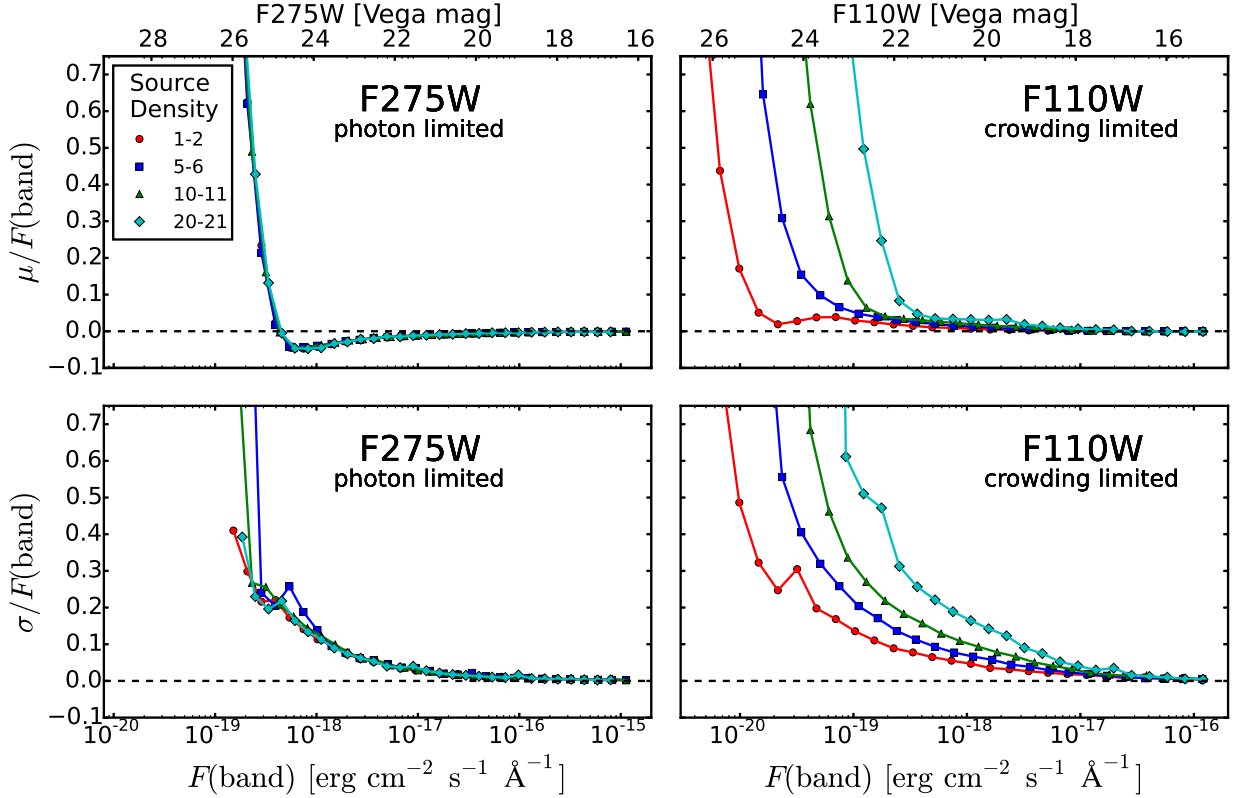
The ASTs used in construction of this noise model were not done simultaneously in all 6 bands, but were instead done in the pairs of bands associated with each camera due to computational resource limitations. Thus, we have conservatively assumed the AST results are independent between all 6 PHAT bands. While we have shown that a noise model built from the full 6 band ASTs allows the BEAST to provide the best constraints on the model parameters (§4.4.2 and 4.4.3), for the remainder of this study we use the noise models built from the single camera ASTs to show an example of the BEAST applied to a large catalog. While the single camera ASTs do not contain sufficient information for full covariance measurements, they do provide good measurements of the dependence of the flux bias and uncertainty in all 6 bands for the full range of fluxes and crowding found in the PHAT survey.

The assumption of band independent ASTs will lead us to overestimate the fit parameter uncertainties (wider

pPDFs) as assuming independent band measurements (no covariance) provides the loosest restriction on the allowed parameter space by the observed SED (see Sec. 4.4.3). In the future, new 6 band ASTs will be computed that will provide the BEAST with a noise model that includes the important covariance between bands due to crowding. In all the following figures, we use the fit parameter expectation values (i.e., pPDF weighted averages) as these reflect the full range of allowed models for each source. This is in contrast to the more usual use of best fit parameter values that only sample the peak of the probability distribution function.

### 5.1. Parameter Sensitivities

We tested how well we can recover the model parameters using the PHAT survey observations with sensitivity tests. The results of sensitivity tests are closely coupled to the assumed noise model. We have chosen to use the noise model from PHAT for Brick 21 to provide realistic sensitivity results using the full PHAT observational strategy, data reduction, and source extraction. For these tests, almost one million models were picked from the set of all PHAT BEAST grid models with F475W Vega magnitudes brighter than 27.6 mag. The distribution of sources were picked by randomly sampling the full prior distribution on the BEAST grid (§4.3) modified by F475W mag limit. This provides a reasonable star formation history, distribution of dust parameters, and PHAT survey sensitivities. Noise was added to these



**Figure 12.** The bias ( $\mu$ ) and uncertainty ( $\sigma$ ) terms for the F275W and F110W filters are plotted. These results were calculated from ASTs performed over the PHAT survey region. The source densities are in units of stars per  $5'' \times 5''$  pixels using the PHAT sources detected in four bands with F475W magnitudes between 24.5 and 27. This range was chosen as the PHAT survey is complete for these brightnesses. The F275W results show the signature of photon limited observations where the behavior is independent of source density. The F110W results are strongly dependent on source density which is the signature of crowding limited observations.

models using the source density dependent PHAT noise model (§5). Thus, these sensitivity tests are directly applicable to the expected BEAST results for real sources in Brick 21. The resulting recovery of models parameters is graphically shown in Fig. 13.

These tests clearly separate the parameters into primary ones that drive the overall SED shape and secondary ones that provide smaller modifications to the overall shape. The primary parameters are  $A(V)$ ,  $\log(t)$ , and  $\log(M)$  and systematically recovered well, except for sources that have  $\log(t) \geq 9$  and  $\log(M) \leq 0.3$ . This systematic error in recovery can be traced to the strong degeneracies between RGB and Red Clump stars for all masses between 1 and  $2 M_{\odot}$ . The random uncertainty in the recovered parameters is approximately 0.5 mag, 0.5, and 0.2 for  $A(V)$ ,  $\log(t)$ , and  $\log(M)$ , respectively. These are averages over the entire parameter range and specific ranges can be better recovered (e.g., low  $A(V)$  values are better recovered than the average).

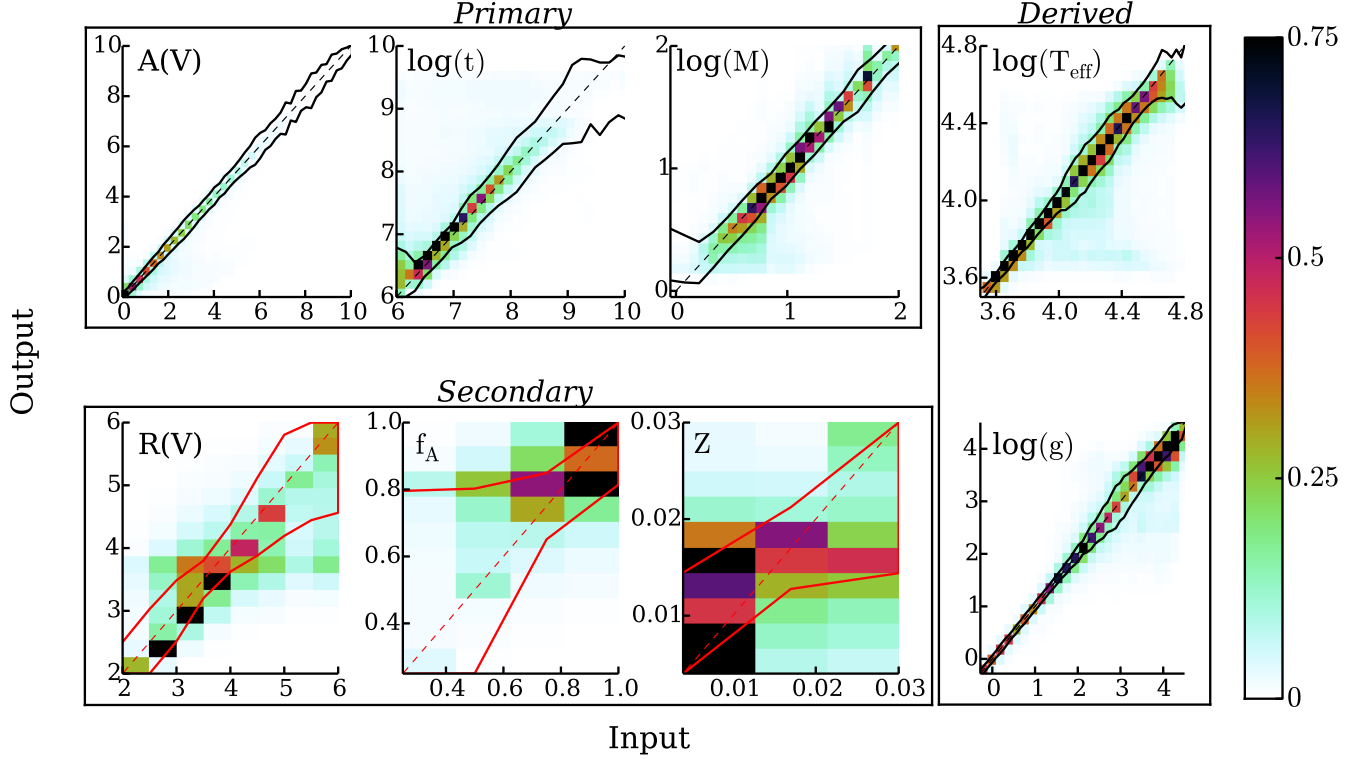
The secondary parameters are  $R(V)$ ,  $f_A$ , and  $Z$ . They are recovered less well than the primary parameters as can be clearly seen from Fig. 13. These secondary parameters modify details of the SED and are more strongly affected by noise than the primary parameters. Of the secondary parameters,  $R(V)$  is recovered the best with mild systematics at high  $R(V)$  values and an average random uncertainty of approximately 1.0. The recovery of  $f_A$  and  $Z$  is fairly poor with systematic offsets and large

random uncertainties of approximately 0.25 and 0.013 for  $f_A$  and  $Z$ , respectively.

The recovery of the derived parameters  $\log(T_{\text{eff}})$  and  $\log(g)$  is quite good and is similar to the primary parameters. These parameters are labeled as derived as they can be directly derived from the primary parameters  $\log(t)$  and  $\log(M)$  at a given  $Z$ . We include the derived parameters here as they are more often used in SED modeling than  $\log(t)$  and  $\log(M)$  and are potentially easier to use to answer specific science questions (e.g., investigating the UV radiation field, Kapala et al., in prep.). They are recovered well systematically over their entire range. The random uncertainties are approximately 0.1 and 0.25 for  $\log(T_{\text{eff}})$  and  $\log(g)$ , respectively. These results show that the BEAST results are not strongly affected by the usually seen degeneracy between  $\log(T_{\text{eff}})$  and  $A(V)$ . This is due to the wide wavelength coverage of the PHAT observations *and* the known distance to the stars allowing the BEAST to fit luminosities instead of fluxes.

## 5.2. Example SED Fit

Fig. 14 gives an example of the BEAST fit to a single source in PHAT Brick 21. This particular source was picked as it has all 6 bands detected with positive flux and represents a common source found in the PHAT survey (i.e., a red clump star). The components of the best fitting models are shown illustrating the strong impact of the photometric measurement bias on the observed flux.



**Figure 13.** The results of sensitivity tests for sources sampled from the model grid with the PHAT noise model are shown as normalized density plots with a log scaling. The contours give the  $1\sigma$  uncertainty (67%) regions illustrating that the primary and derived parameters are well recovered overall. The precision of the recovery of the secondary parameters is lower than for the primary parameters as they are more strongly influenced by small changes in the SED shapes. The density normalization is for display purposes. The distribution of sources is a reflection of the recovery and the density of simulated sources which itself is a function of the model priors and PHAT observation sensitivity.

The 1D pPDFs show a range of behaviors. The  $\log(M)$ ,  $\log(t)$ ,  $R(V)$ ,  $\log(T_{\text{eff}})$ , and  $\log(g)$  show single peaked pPDFs. The  $A(V)$  pPDF shows a double peaked pPDFs. Finally, the  $f_A$  and  $Z$  pPDFs show sloped pPDFs with no clear peak. This wide range of non-Gaussian and multi-peaked pPDFs illustrates the complexity of this type of SED fitting and the importance of fully mapping the pPDF. This complexity is one of the primary motivations for basing the BEAST on a grid fitting technique.

### 5.3. Recovered Hertzsprung-Russel Diagram

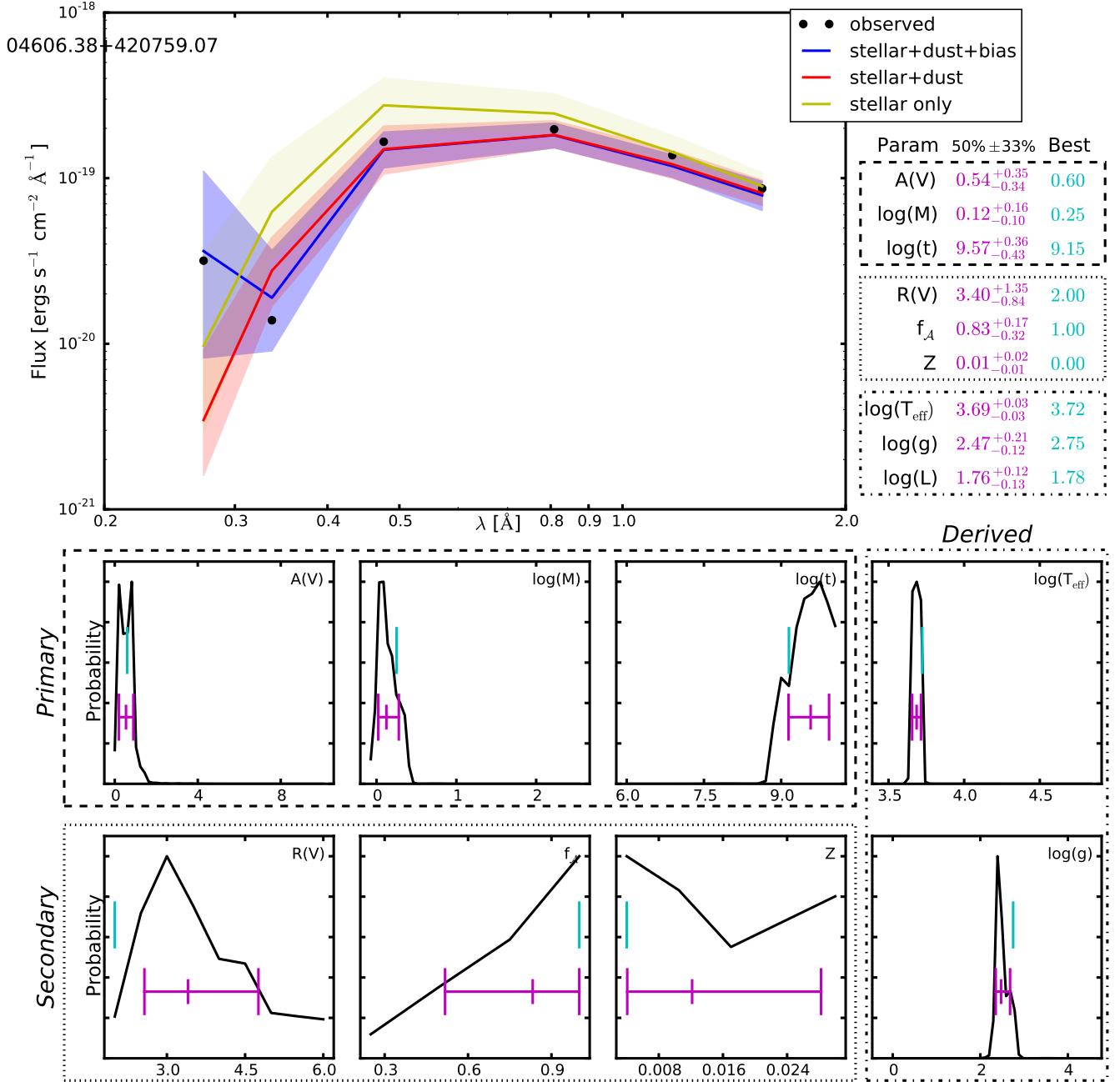
The Hertzsprung-Russel diagram for all the Brick 21 sources fit is given in Fig. 15. The distribution of sources is reasonable, with sources on the RGB dominating. The main sequence is well populated, consistent with the large number of star forming regions present in Brick 21. The lower region shows the sensitivity limit that is the result of requiring detections in at least four bands coupled with the varying dust columns and crowding noise across this brick. Comparing this figure with Fig. 8 indicates that this region has, not surprisingly, a somewhat different star formation history than we assumed in our priors. This difference provides confirmation that the BEAST results are not dominated by the priors.

### 5.4. Fit Uncertainties

For all the 0.7 million sources fitted in Brick 21, Figs. 16, 17, & 18 show the primary, secondary, and de-

rived  $1\sigma$  fit parameter uncertainties versus their expectation values. In addition these figures give the average fit parameter uncertainty color coded on HR diagrams. The  $1\sigma$  uncertainties are measured from the 67% width of the 1D pPDFs. These figures illustrate the precision that is possible with the BEAST for the PHAT observations of the Brick 21 region for sources detected in at least four bands with the assumed physical and noise models. These figures provide a detailed view how the fit parameter precision versus fit parameter and location in the HR diagram. The fit parameter accuracy is discussed in §5.1 as this can only be tested with simulations. The precision in both the real and simulated data for this Brick are similar.

Fig. 16 clearly illustrates that the precision in the recovery of the primary parameters is good, but variable. For example, the  $A(V)$  plots show a bimodal precision with concentrations at  $\sim 0.4$  and  $\sim 1.0$  mag. The higher precision results can be seen via the HR diagram to be the result of sources on the main sequence and red giant branch. The low precision results are due to sources at the tip of the RGB as well as the faintest sources. The results for the RGB tip sources is likely due to the lack of Asymptotic Giant Branch stars in the BEAST model as they are missing from the PARSEC isochrones used. The results for the faint sources are simply the result of low signal-to-noise measurements. The results for the other two primary parameters are similar in that the RGB tip



**Figure 14.** The fitting results for an individual star in Brick 21 are shown. While this figure illustrates the fit to an individual source, it is for a red giant star that is one of the common sources we find in the PHAT region. The 50% model is shown along with shaded colored regions indicating the range of models that fit within  $1\sigma$ . The models shown are the full model including the observational bias (stellar+dust+bias), the physical model alone (stellar+dust), and the stellar model alone (stellar). The impact of the bias in the photometric measurements is clearly seen in the UV and illustrates the importance of a full physical and observational model for this type of fitting. The best fit parameter values (cyan) and  $50\% \pm 33\%$  values (purple) are given numerically as well as graphically in the 1D pPDF plots. The 1D pPDF plots give the normalized probabilities with the y-axis scale ranging from 0 to 1.1. The best fit values are determined from the full 6D pPDFs and, as such, often do not appear at the peak of the 1D pPDFs.

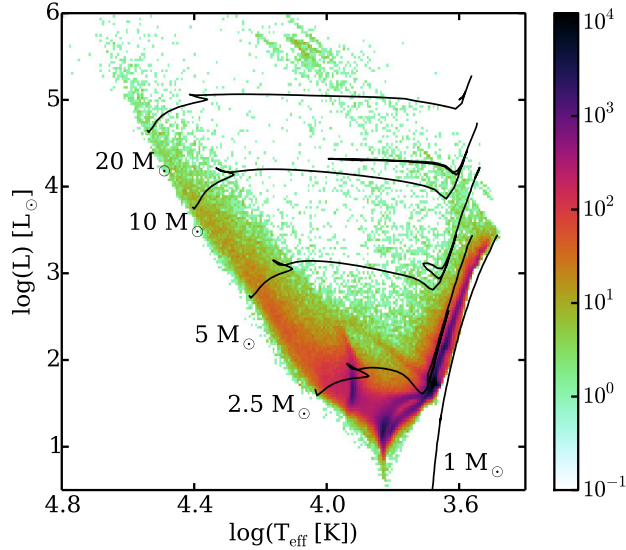
sources and faint sources are less well recovered. But there are differences. The  $\log(M)$  plots show that the mass is well recovered on the main sequence, and less well recovered on the red giant branch. The combined low precision results for  $\log(M) < 0.3$  and  $\log(t) > 9$  can be traced to RGB and Red Clump stars being very degenerate in the HR diagram for all masses between 1 and  $2 M_{\odot}$  at the precision of the PHAT observations.

The precision in the recovery of the secondary param-

eters shown in Fig. 17 is fairly well behaved with a fairly clear relationship between parameter and uncertainty. The HR diagram for  $R(V)$  shows a strong gradient from bright to faint sources, illustrating that the  $R(V)$  precision is driven by the signal-to-noise in the observations. The  $f_A$  and  $Z$  HR diagrams show fairly low precisions throughout indicating the overall difficulty in recovering these two parameters given the PHAT bands and depth.

The precision of the recovery of the derived param-





**Figure 15.** Using the Brick 21 results, the density of sources is shown in a Hertzsprung Russell diagram. The overall structure is reasonable, but there are nonphysical streaks in the upper portion and ridges in the lower portion. The nonphysical ridges in the faint source population are due to the sensitivity limits of the PHAT survey sources used. We required each source to be detected in four PHAT bands and to have a F475W magnitude above 27.6. In general, the particular set of four PHAT bands fulfilling this requirement varies from being UV/optical to optical/NIR and this is the expected origin of the two prominent ridges in the faint source population. The nonphysical streaks for the brighter sources are likely due to variable survey depth due to gaps between the chips in WFC3 and ACS and/or an issue with the adopted BEAST grid being too coarse in this region of the HR diagram.

ters (Fig. 18) show behaviors similar to those seen for the primary parameters. The main sequence and RGB stars have better precision than the RGB tip and the fainter sources. The behavior for lower precision for sources with  $\log(T_{\text{eff}})$  values between 3.75 and 3.9 is a result of residual degeneracies between  $T_{\text{eff}}$  and  $A(V)$  (Romaniello et al. 2002b).

It is worth remembering that these precisions are based on a noise model that does not take into account any covariance. We have shown in §4.4.3 that once we have ASTs measured simultaneously in all 6 bands, we will be able to improve the precision of the BEAST fits and potentially reduce the degeneracies between parameters.

### 5.5. Dust Maps

One of the specific goals of the BEAST effort is to study the dust properties in galaxies. A map of the dust column density can be derived by averaging the results for individual stars in rectangular regions. This gives an estimate of the average dust column density in these regions after multiplication by a factor of two as this corrects for the fact that (on average) half of the stars will be behind the dust and the other half in front in a disk galaxy. This was done for the PHAT Brick 21 sources and the resulting  $A(V)$  map is given in Fig. 19. There are coherent structures in the  $A(V)$  map and similar structures are also seen in the dust surface density map derived from fitting the dust emission infrared maps (Draine et al. 2014). The variations between these two images provide information on real variations in the IR dust emissivity modulo systematics in the BEAST de-

rived  $A(V)$  map (e.g., sensitivity to the highest columns of dust). In future work, we will derive a more accurate  $A(V)$  map using a detailed dust and star geometry model, the full pPDFs, and accounting for the completeness of the PHAT observations (Arab et al., in prep.). The comparison in this paper is presented to show that the BEAST produces reasonable results for dust column densities.

## 6. SUMMARY

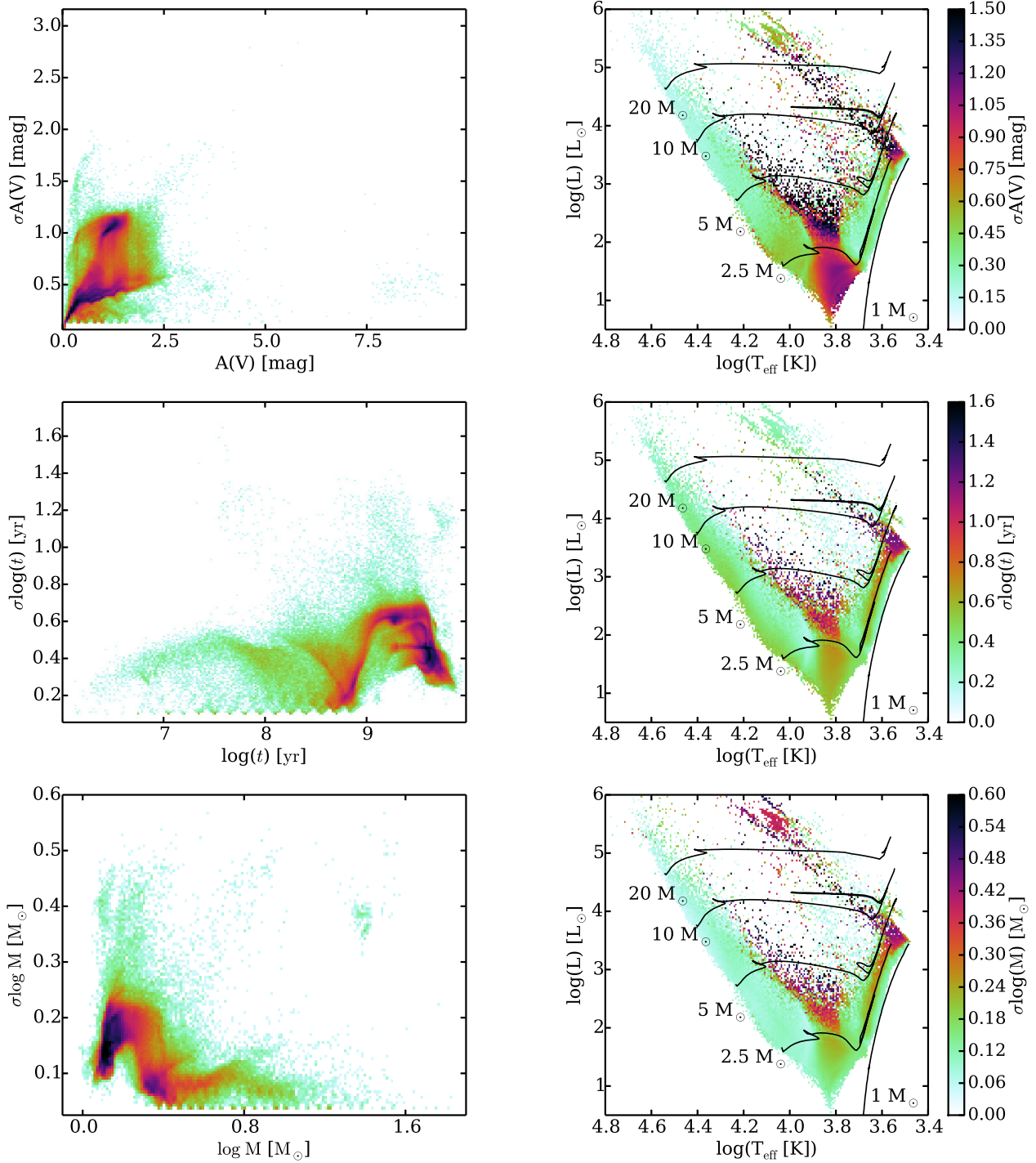
We have presented the BEAST, a probabilistic approach to modeling the photometric SEDs of sources detected in large resolved star surveys. The BEAST uses a 6 parameter model of the SED of individual stars extinguished by dust. The stellar portion of the model is based on stellar evolution and atmosphere models and has the parameters of age ( $t$ ), mass ( $M$ ), and metallicity ( $Z$ ). The dust extinction portion of the model is based on a newly developed mixture model of dust extinction and has the parameters of dust column ( $A(V)$ ), average dust grain size ( $R(V)$ ), and mixture coefficient ( $f_A$ ). This mixture model has the feature of encompassing all observed extinction curves in the Milky Way, Large Magellanic Cloud, and Small Magellanic Cloud.

The BEAST fitting technique allows for bias and covariance between bands in the observed SEDs using a multi-variate Normal/Gaussian distribution. The noise model used includes terms for the photon, crowding, and absolute calibration noise. The importance of correctly integrating the model SEDs using the full photometric band response functions was illustrated. The impact on the fit parameter precision and accuracy of correctly accounting for the covariance between observed bands in the noise model was shown. Overall, the strength of the BEAST over most existing SED fitting codes is a careful treatment of biases and correlations in the uncertainties between bands and the new dust extinction mixture model.

We illustrated the application of the BEAST to real data using the PHAT survey of M31. The assumed priors are based on independent knowledge of stellar and dust physics and were presented in the context of fitting using a grid. The derivation of the noise model from ASTs and the details of the HST absolute flux calibration were described. Using sensitivity tests, we illustrated that overall the BEAST does a good job recovering the input fit parameters. Finally, we presented the results from fitting all the sources detected in 4 bands in Brick 21 of the PHAT survey. These results yield a reasonable HR diagram, understandable variations in fit parameter precision, and a spatial map of the average dust column that correlates well with existing infrared derived dust surface densities.

While the BEAST has been motivated and will be used for the PHAT survey to study the stellar and dust content of M31, it can be used to fit the data from any resolved star extragalactic survey.

We thank the referee for an insightful comments that motivated us to significantly improve the content and presentation of this paper. Support for program # 12055 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc.,



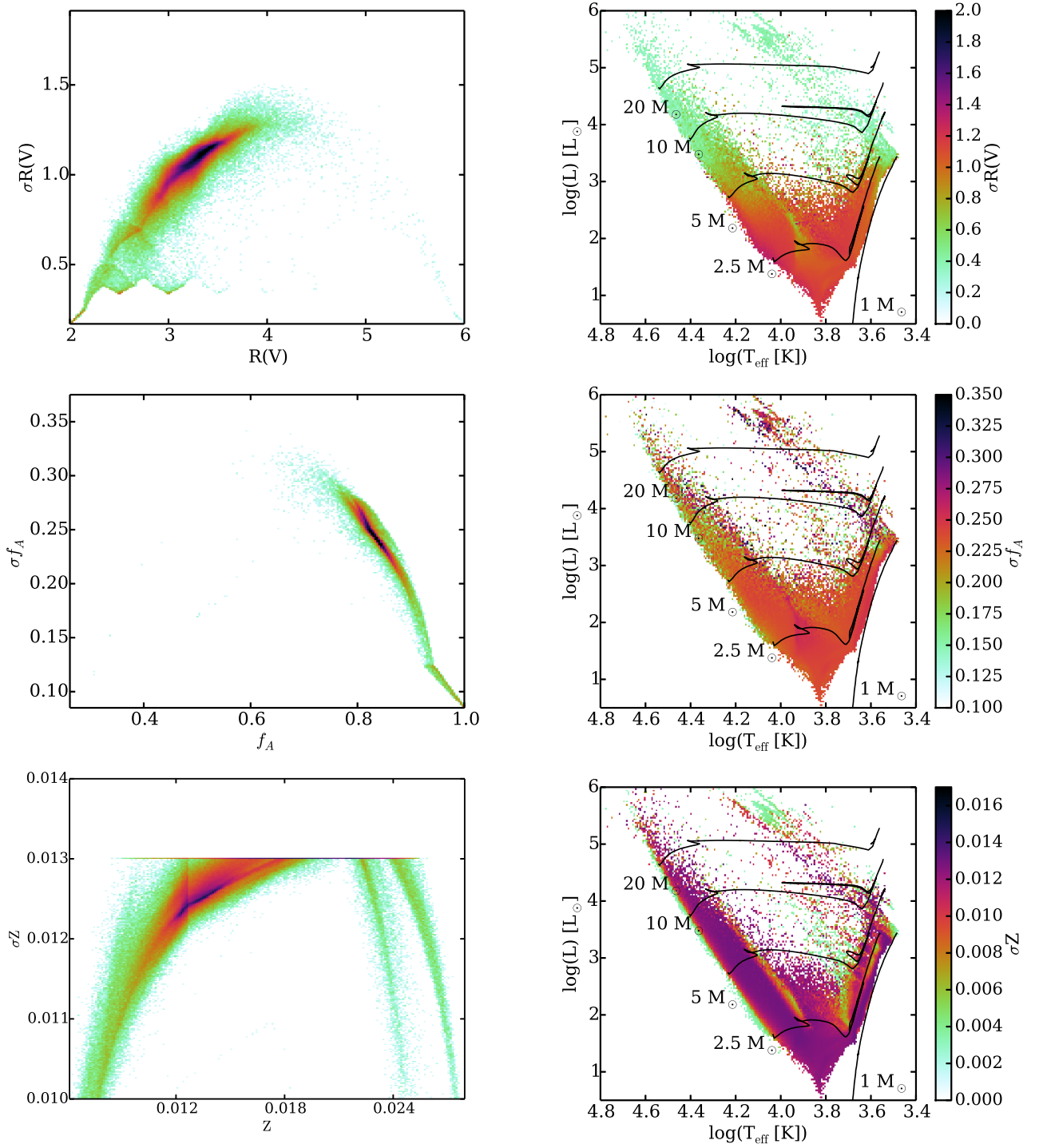
**Figure 16.** In the left column, the  $1\sigma$  uncertainties are plotted versus the expectation values for the primary fit parameters color coded by log density (green low, black high) from the Brick 21 fitting. In the right column, the average  $1\sigma$  uncertainty at each position in a HR diagram is shown. The careful reader will notice that the right  $t$  plot looks like the Loch Ness Monster.

under NASA contract NAS 5-26555. Support for DRW is provided by NASA through Hubble Fellowship grants HST-HF-51331.01 awarded by the Space Telescope Science Institute. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575. The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing High Performance

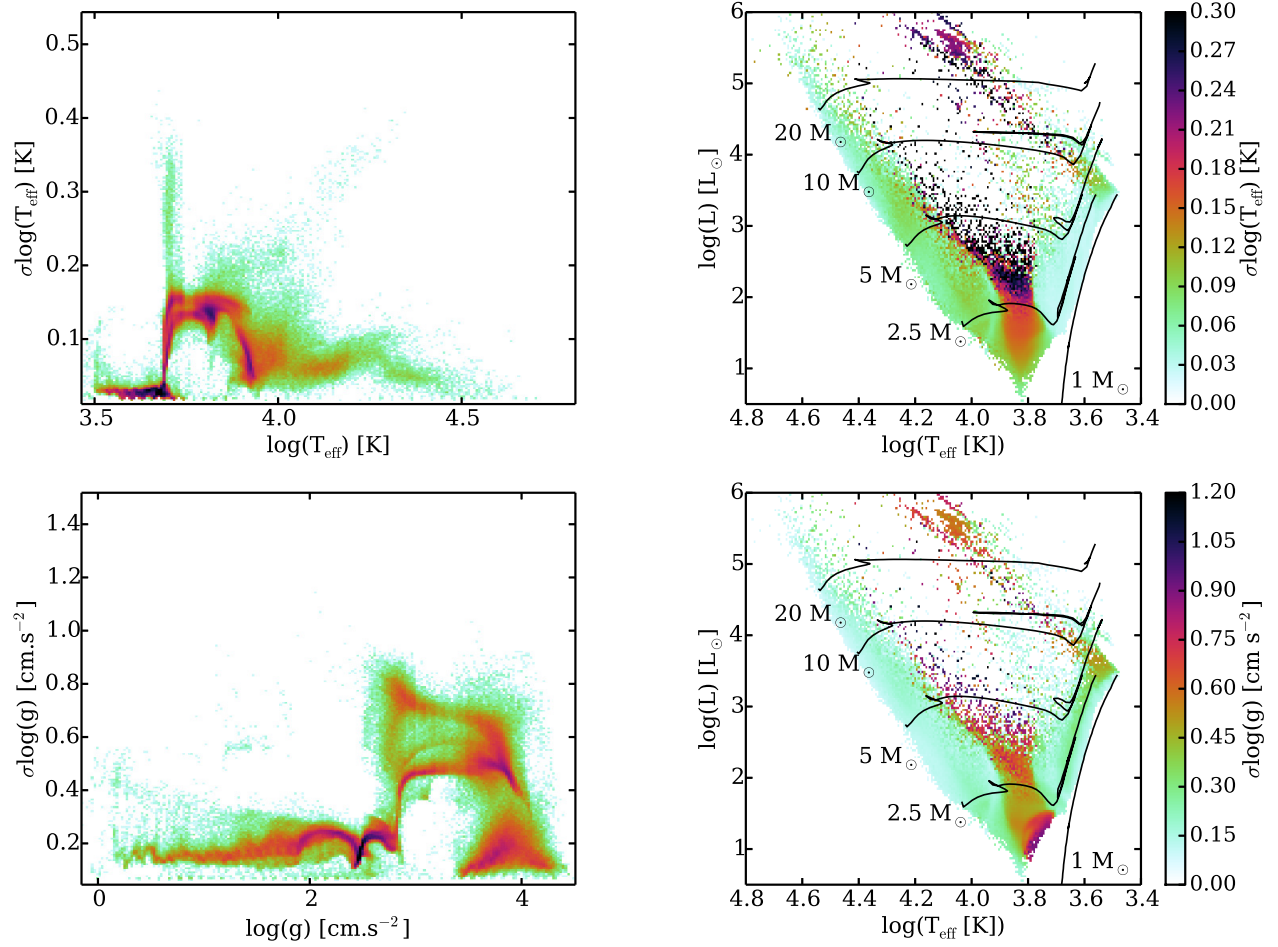
Computing (HPC) resources that have contributed to the research results reported within this paper. URL: <http://www.tacc.utexas.edu>.

#### REFERENCES

- Anderson, J., King, I. R., Richer, H. B., et al. 2008, *AJ*, 135, 2114
- Bailer-Jones, C. A. L. 2011, *MNRAS*, 411, 435
- Bastian, N., Covey, K. R., & Meyer, M. R. 2010, *ARA&A*, 48, 339

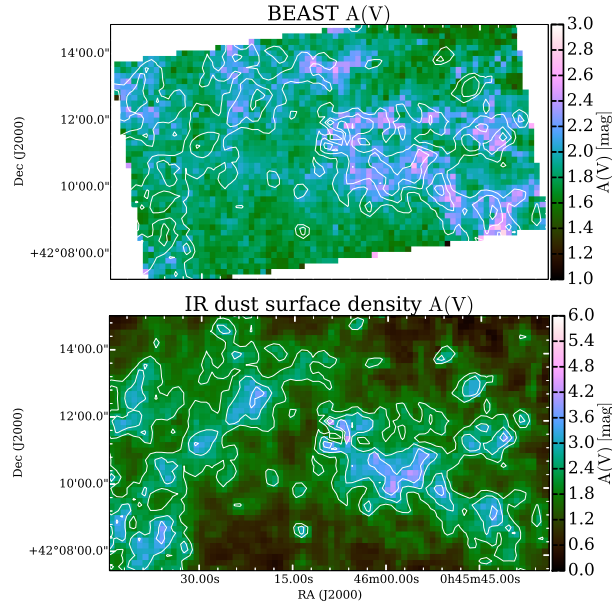


**Figure 17.** In the left column, the  $1\sigma$  uncertainties are plotted versus the expectation values for the secondary fit parameters color coded by log density (green low, black high) from the Brick 21 fitting. In the right column, the average  $1\sigma$  uncertainty at each position in a HR diagram is shown. The sharp upper boundary in  $\sigma Z$  is the when the pPDF is completely unconstrained and is one half the allowed range in  $Z$ .



**Figure 18.** In the left column, the  $1\sigma$  uncertainties are plotted versus the expectation values for the derived fit parameters color coded by log density (green low, black high) from the Brick 21 fitting. In the right column, the average  $1\sigma$  uncertainty at each position in a HR diagram is shown.





**Figure 19.** The  $A(V)$  map derived from averaging the BEAST results for PHAT brick 21 (top) and the IR derived dust surface density (bottom, Draine et al. 2014) are shown. The contours in both panels show the IR derived dust surface density. The overall scale is 2x higher in the IR derived dust map as the Draine et al. (2014) prediction of  $A(V)$  is known to be systematically 2x higher due to dust grain model calibration issues (Lombardi et al. 2014; Dalcanton et al. 2015; Planck Collaboration et al. 2016). The two panels show very similar morphologies.

Bergemann, M., Kudritzki, R.-P., Plez, B., et al. 2012, *ApJ*, 751, 156  
 Berry, M., Ivezić, Ž., Sesar, B., et al. 2012, *ApJ*, 757, 166  
 Bianchi, L., Clayton, G. C., Bohlin, R. C., Hutchings, J. B., & Massey, P. 1996, *ApJ*, 471, 203  
 Bianchi, L., Efremova, B., Hodge, P., & Kang, Y. 2012a, *AJ*, 144, 142  
 Bianchi, L., Efremova, B., Hodge, P., Massey, P., & Olsen, K. A. G. 2012b, *AJ*, 143, 74  
 Bianchi, L., Scuderi, S., Massey, P., & Romaniello, M. 2001, *AJ*, 121, 2020  
 Bohlin, R. C., Gordon, K. D., & Tremblay, P.-E. 2014, *ArXiv e-prints*, arXiv:1406.1707  
 Bressan, A., Marigo, P., Girardi, L., Nanni, A., & Rubele, S. 2013, in *European Physical Journal Web of Conferences*, Vol. 43, European Physical Journal Web of Conferences, 3001  
 Bressan, A., Marigo, P., Girardi, L., et al. 2012, *MNRAS*, 427, 127  
 Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, 345, 245  
 Castelli, F., & Kurucz, R. L. 2004, *ArXiv Astrophysics e-prints*, arXiv:astro-ph/0405087  
 Chen, Y., Girardi, L., Bressan, A., et al. 2014, *MNRAS*, 444, 2525  
 Da Rio, N., & Robberto, M. 2012, *AJ*, 144, 176  
 Dalcanton, J. J., Williams, B. F., Seth, A. C., et al. 2009, *ApJS*, 183, 67  
 Dalcanton, J. J., Williams, B. F., Melbourne, J. L., et al. 2012a, *ApJS*, 198, 6  
 Dalcanton, J. J., Williams, B. F., Lang, D., et al. 2012b, *ApJS*, 200, 18

—. 2012c, *ApJS*, 200, 18  
 Dalcanton, J. J., Funesneau, M., Hogg, D. W., et al. 2015, *ApJ*, 814, 3  
 Dolphin, A. E. 2000, *PASP*, 112, 1383  
 Draine, B. T., Aniano, G., Krause, O., et al. 2014, *ApJ*, 780, 172  
 Fitzpatrick, E. L. 1999, *PASP*, 111, 63  
 Funesneau, M., & Lançon, A. 2010, *A&A*, 521, A22  
 Girardi, L., Williams, B. F., Gilbert, K. M., et al. 2010, *ApJ*, 724, 1030  
 Gordon, K. D., Cartledge, S., & Clayton, G. C. 2009, *ApJ*, 705, 1320  
 Gordon, K. D., & Clayton, G. C. 1998, *ApJ*, 500, 816  
 Gordon, K. D., Clayton, G. C., Misselt, K. A., Landolt, A. U., & Wolff, M. J. 2003, *ApJ*, 594, 279  
 Green, D. A. 2011, *Bulletin of the Astronomical Society of India*, 39, 289  
 Green, G. M., Schlafly, E. F., Finkbeiner, D. P., et al. 2014, *ApJ*, 783, 114  
 Hogg, D. W., Baldry, I. K., Blanton, M. R., & Eisenstein, D. J. 2002, *ArXiv Astrophysics e-prints*, astro-ph/0210394  
 Holtzman, J. A., Afonso, C., & Dolphin, A. 2006, *ApJS*, 166, 534  
 Johnson, H. L., & Morgan, W. W. 1953, *ApJ*, 117, 313  
 Kroupa, P. 2001, *MNRAS*, 322, 231  
 Lanz, T., & Hubeny, I. 2003, *ApJS*, 146, 417  
 —. 2007, *ApJS*, 169, 83  
 Lombardi, M., Bouy, H., Alves, J., & Lada, C. J. 2014, *A&A*, 566, A45  
 Maíz-Apellániz, J. 2004, *PASP*, 116, 859  
 Maíz Apellániz, J., & Rubio, M. 2012, *A&A*, 541, A54  
 Marigo, P., Girardi, L., Bressan, A., et al. 2008, *A&A*, 482, 883  
 Mathis, J. S., & Cardelli, J. A. 1992, *ApJ*, 398, 610  
 McConnell, A. W., Irwin, M. J., Ibata, R. A., et al. 2009, *Nature*, 461, 66  
 Misselt, K. A., Clayton, G. C., & Gordon, K. D. 1999, *ApJ*, 515, 128  
 Ness, M., Hogg, D. W., Rix, H.-W., Ho, A. Y. Q., & Zasowski, G. 2015, *ApJ*, 808, 16  
 Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, 586, A132  
 Rauch, T., Werner, K., Bohlin, R., & Kruk, J. W. 2013, *A&A*, 560, A106  
 Romaniello, M., Panagia, N., Scuderi, S., & Kirshner, R. P. 2002a, *AJ*, 123, 915  
 —. 2002b, *AJ*, 123, 915  
 Sabbi, E., Anderson, J., Lennon, D. J., et al. 2013, *AJ*, 146, 53  
 Schlafly, E. F., & Finkbeiner, D. P. 2011, *ApJ*, 737, 103  
 Schönrich, R., & Bergemann, M. 2014, *MNRAS*, 443, 698  
 Sirianni, M., Jee, M. J., Benítez, N., et al. 2005, *PASP*, 117, 1049  
 Stetson, P. B., & Harris, W. E. 1988, *AJ*, 96, 909  
 Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, *AJ*, 123, 485  
 Tang, J., Bressan, A., Rosenfield, P., et al. 2014, *MNRAS*, 445, 4287  
 Taylor, J. 1997, *Introduction to Error Analysis, the Study of Uncertainties in Physical Measurements*, 2nd Edition (University Science Books)  
 Towns, J., Cockerill, T., Dahan, M., et al. 2014, *Computing in Science and Engineering*, 16, 62  
 Valencic, L. A., Clayton, G. C., & Gordon, K. D. 2004, *ApJ*, 616, 912  
 Wallerstein, G., & Carlson, M. 1960, *ApJ*, 132, 276  
 Weisz, D. R., Johnson, L. C., Foreman-Mackey, D., et al. 2015, *ApJ*, 806, 198  
 Westera, P., Lejeune, T., Buser, R., Cuisinier, F., & Bruzual, G. 2002, *A&A*, 381, 524  
 Williams, B. F., Lang, D., Dalcanton, J. J., et al. 2014, *ArXiv e-prints*, arXiv:1409.0899  
 Zaritsky, D., Harris, J., & Thompson, I. 1997, *AJ*, 114, 1002  
 Zaritsky, D., Harris, J., Thompson, I. B., & Grebel, E. K. 2004, *AJ*, 128, 1606